

# SEGMENTASI DAERAH RENTAN STUNTING DI INDONESIA MENGGUNAKAN METODE K-MEANS

Eko Saputra<sup>1</sup>, Hardian Oktavianto<sup>2</sup>, Lutfi Ali Muharom<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

E-mail: eko.syahputra543@gmail.com<sup>1</sup>, hardian@unmuahjember.ac.id<sup>2</sup>,

lutfi.muharom@unmuahjember.ac.id<sup>3</sup>

## ABSTRAK

Saat ini, Indonesia merupakan salah satu negara dengan *prevalensi stunting* yang cukup tinggi dibandingkan dengan negara-negara berpendapatan menengah lainnya. Dalam narasi Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024 penekanan angka *stunting* ditargetkan menjadi 19% pada 2024. Agar dapat diketahui bagaimana perkembangan daerah-daerah di Indonesia dalam penanganan *stunting* berdasarkan data persentase pemberian asi eksklusif, cakupan imunisasi dasar lengkap serta status gizi pada balita sejak tahun 2016 sampai tahun 2018, maka dalam penelitian ini membahas mengenai pengelompokan provinsi di Indonesia yang rentan terjadi *stunting* menggunakan metode *K-Means* dengan menghitung nilai *Sum of Squares Error (SSE)* dalam menentukan *cluster* terbaiknya. Hasil yang diperoleh dengan skenario 2 *cluster* sampai 6 *cluster* yaitu *cluster* terbaik berada pada 6 *cluster* dengan nilai *Sum of Squares Error (SSE)* sebesar 12015.66863, sedangkan jumlah anggota pada masing-masing *cluster* yaitu *cluster* 1 terdapat 18 provinsi, *cluster* 2 terdapat 2 provinsi, *cluster* 3 terdapat 6 provinsi, *cluster* 4 terdapat 3 provinsi, *cluster* 5 terdapat 1 provinsi dan *cluster* 6 terdapat 4 provinsi. Kemudian untuk kelompok provinsi pada *cluster* 1 memiliki rata-rata persentase paling tinggi pada atribut ASI dan IDL serta paling rendah untuk atribut GB, SK dan SP sehingga resiko terjadinya *stunting* sangat rendah. Sedangkan kelompok provinsi yang sangat rentan terjadi *stunting* berada pada *cluster* 5 dengan rata-rata persentase terendah pada atribut ASI dan IDL serta persentase tertinggi untuk atribut GB, SK dan SP.

**Kata kunci:** *Stunting, Cluster, K-Means, Sum of Squares Error (SSE)*.

## ABSTRACT

Currently, Indonesia is one of the countries with a fairly high prevalence of stunting compared to other middle-income countries. In the narrative of the National Medium-Term Development Plan (RPJMN) 2020-2024, the stunting rate is targeted to be 19% by 2024. In order to know how the development of regions in Indonesia in handling stunting is based on data on the percentage of exclusive breastfeeding, complete basic immunization coverage and nutritional status in toddlers from 2016 to 2018, this study discusses the grouping of provinces in Indonesia that are prone to stunting using the *K-Means* method by calculating the *Sum of Squares Error (SSE)* value in determining the best cluster. The results obtained with the scenario 2 clusters to 6 clusters, namely the best clusters are in 6 clusters with a *Sum of Squares Error (SSE)* value of 12015.66863, while the number of members in each cluster, namely cluster 1, has 18 provinces, cluster 2 has 2 provinces, cluster 3 has 6 provinces, cluster 4 has 3 provinces, cluster 5 has 1 province and cluster 6 has 4 provinces. Then for the provincial group in cluster 1 has the highest average percentage on attributes of ASI and IDL and the lowest for attributes of GB, SK and SP so that the risk of stunting is very low. Meanwhile, the province group which is very prone to stunting is in cluster 5 with the lowest average percentage for attributes of ASI and IDL and the highest percentage for attributes of GB, SK and SP.

**Keywords:** *Stunting, Cluster, K-Means, Sum of Squares Error (SSE)*.

## PENDAHULUAN

### Latar Belakang

Saat ini, Indonesia merupakan salah satu negara dengan *prevalensi stunting* yang cukup tinggi dibandingkan dengan negara-negara berpendapatan menengah lainnya. Situasi ini jika tidak diatasi dapat mempengaruhi kinerja pembangunan Indonesia baik yang menyangkut pertumbuhan ekonomi, kemiskinan dan ketimpangan (Tim Nasional Percepatan Penanggulangan Kemiskinan, Sekretariat Wakil Presiden, 2017).

Kementerian kesehatan bekerja sama dengan Badan Pusat Statistik (BPS) secara periodik 5 tahunan melakukan riset yang dilakukan terhadap 84.000 balita dalam bentuk Hasil Studi Status Gizi Balita Indonesia (SSGBI). Hasilnya *prevalensi* balita *underweight* atau gizi kurang pada 2019 berada di angka 16,29 persen. Angka ini mengalami penurunan sebanyak 1,5 persen. Kemudian *prevalensi* balita *stunting* pada 2019 sebanyak 27,67 persen, turun sebanyak 3,1 persen. Sementara itu untuk *prevalensi* balita *wasting* (kurus), berada pada angka 7,44 persen. Angka ini turun 2,8 persen. Semua data dibandingkan dengan hasil survei dari tahun 2018. Dalam narasi Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024 penekanan angka *stunting* ditargetkan menjadi 19% pada 2024 (Riskesdas, 2018).

Agar dapat diketahui bagaimana perkembangan daerah-daerah di Indonesia dalam penanganan *stunting*, maka perlu adanya pengelompokan daerah, dalam hal ini daerah yang dikelompokkan meliputi 34 provinsi di Indonesia. Penggunaan teknik *Data Mining* dapat menjadi alternatif dalam menentukan kelompok-kelompok wilayah yang memiliki kesamaan karakter data pada setiap wilayah. Diantara algoritma *clustering* lainnya, algoritma *K-Means* paling sering digunakan dan baik dalam algoritma *Partitional Clustering* karena sederhana dan lebih efisien (Budiman, 2012).

### Rumusan Masalah

Permasalahan yang dapat dirumuskan berkaitan dengan *segmentasi* daerah rentan *stunting* di Indonesia adalah: (1) Berapa jumlah *cluster* optimum ditinjau dari persentase status gizi dan pemberian ASI ekslusif serta cakupan imunisasi dasar lengkap pada balita menggunakan metode algoritma *K-Means*, (2) Bagaimana kelompok wilayah provinsi di Indonesia yang terdapat dalam jumlah *cluster* yang optimum.

### Tujuan Penelitian

Tujuan penelitian yang hendak dicapai adalah: (1) Mengetahui keberhasilan penggunaan algoritma *K-Means* pada pengelompokan wilayah provinsi di Indonesia berdasarkan persentase status gizi dan pemberian ASI ekslusif serta cakupan imunisasi dasar lengkap pada balita. (2) Mencari jumlah *cluster* optimum ditinjau dari persentase status gizi dan pemberian ASI ekslusif serta cakupan imunisasi dasar lengkap pada balita menggunakan metode algoritma *K-Means*. (3) Mengetahui kelompok wilayah provinsi di Indonesia yang rentan terjadi *Stunting* yang terdapat dalam satu *cluster*.

### Manfaat Penelitian

Manfaat yang dapat diperoleh dari penyusunan tugas akhir ini adalah: Bagi pemerintah yaitu memberikan pengetahuan berupa kelompok wilayah di Indonesia yang rentan kejadian *stunting* dan sebagai alternatif untuk menilai perkembangan daerah dalam penanganan *stunting*. Bagi Peneliti yaitu mengetahui *cluster* optimum dan kelompok wilayah provinsi yang rentan terjadi *stunting* berdasarkan persentase status gizi dan pemberian ASI ekslusif serta cakupan imunisasi dasar lengkap pada balita. Bagi peneliti lain sebagai referensi dalam penelitian yang akan datang, khususnya dalam hal pengembangan dan implementasi teknologi data mining.

## METODE PENELITIAN

## Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data persentase status gizi balita (gizi buruk = indeks BB/U, sangat pendek = indeks TB/U dan sangat kurus = indeks BB/TB), persentase cakupan dasar imunisasi lengkap diatas 80% dan persentase pemberian ASI eksklusif pada balita tiap provinsi di Indonesia tahun 2016 sampai tahun 2018. Pengumpulan data utama yang akan diproses dengan algoritma *K-Means* diambil dari *website* kementerian kesehatan Republik Indonesia bagian publikasi data serta jurnal dan situs internet yang berkaitan untuk melengkapi data yang diperlukan. Jenis data yang digunakan dalam penelitian ini adalah data sekunder.

## Implementasi Algoritma K-Means

Adapun proses algoritma *K-Means* dalam pengelompokan faktor penyebab terjadi *stunting* adalah sebagai berikut:

## Iterasi Ke-1

- ### 1. Menentukan jumlah *cluster*

Jumlah *cluster* yang digunakan sebanyak 3 *cluster*

- ## 2. Menentukan nilai pusat *centroid*

Penentuan pusat awal *centroid* ditentukan secara *random/acak* diambil dari data yang ada.

	C1	C2	C3
ASI 2016	12.40	28.91	55.40
IDL 2016	34.48	72.97	100.00
GB 2016	1.30	3.38	6.67
SP 2016	4.66	8.64	12.33
SK 2016	1.30	3.51	9.03
ASI 2017	10.73	32.49	61.45
IDL 2017	24.14	82.96	100.00
GB 2017	2.10	4.51	6.80
SP 2017	5.10	10.82	15.90
SK 2017	1.60	3.82	6.00
ASI 2018	20.43	56.36	78.63
IDL 2018	21.74	71.22	100.00
GB 2018	2.30	4.62	7.40
SP 2018	6.10	11.13	16.00
SK 2018	1.10	3.67	5.00

Tabel 3.1 Titik awal pusat centroid

### 3. Menghitung jarak dari *centroid*

Menghitung jarak antara titik *centroid* dengan titik tiap objek dengan menggunakan persamaan *Euclidean Distance*

Maka akan didapatkan nilai matrik jarak sebagai berikut:

$$C_{1z} = \sqrt{(23.50 - 12.40)^2 + (69.57 - 34.48)^2 + (2.55 - 1.30)^2 + (7.56 - 4.66)^2 + (3.51 - 1.30)^2 + (22.99 - 10.73)^2 + (52.17 - 24.14)^2 + (5.90 - 6.80)^2 + (12.20 - 15.90)^2 + (3.70 - 5.80)^2 + (60.84 - 48.32)^2 + (21.74 - 34.48)^2 + (6.70 - 5.10)^2 + (16.00 - 15.30)^2 + (5.00 - 4.80)^2}$$

$$C_{2z} = \frac{(23.50 - 28.91)^2 + (69.57 - 72.97)^2 + (2.55 - 3.38)^2 + (7.56 - 8.64)^2 + (3.51 - 3.51)^2 + (22.99 - 32.49)^2 + (52.17 - 82.96)^2 + (5.90 - 4.51)^2 + (12.20 - 10.82)^2 + (3.70 - 3.82)^2 + (60.84 - 56.36)^2 + (21.74 - 71.22)^2 + (6.70 - 4.62)^2 + (16.00 - 11.13)^2 + (5.00 - 3.67)^2}{59.85772152}$$

$$C_{3\beta} = \sqrt{\frac{(23.50 - 55.40)^2 + (69.57 - 100.00)^2 + (2.55 - 6.67)^2 + (7.56 - 12.33)^2 + (3.51 - 9.03)^2 + (22.99 - 61.45)^2 + (52.17 - 100.00)^2}{(5.90 - 6.80)^2 + (12.20 - 15.90)^2 + (3.70 - 6.00)^2 + (60.84 - 78.63)^2 + (21.74 - 100.00)^2 + (6.70 - 7.40)^2 + (16.00 - 16.00)^2}} + (5.00 - 5.00)^2$$

Dan seterusnya dilanjutkan menghitung untuk data ke-2 ... N terhadap pusat awal *cluster* hingga didapatkan nilai matriks jarak.

- Pengelompokan objek, jarak hasil perhitungan pada point ke-3 akan dilakukan perbandingan dan dipilih jarak yang terdekat antara data dengan pusat *cluster*, yang akan menentukan bahwa data yang memiliki jarak terdekat akan berada dalam satu kelompok dengan pusat *cluster* terdekat, pengelompokan objek/data tersebut dapat dilihat pada tabel sebagai berikut:

**Tabel 3.2** Perhitungan jarak dan pengelompokan data *iterasi* ke-1

Data ke	Provinsi	C1	C2	C3	KC1	KC2	KC3
1	Aceh	64.34416291	59.85772152	110.6438647		ok	
2	Sumatera Utara	81.88644332	31.04809654	90.49798175		ok	
3	Sumatera Selatan	136.839552	42.91158669	30.47371162		ok	
4	Bengkulu	133.15598	40.71139061	49.99700691		ok	
5	DKI Jakarta	144.7354051	49.88743782	27.94800172		ok	
6	Jawa Barat	136.245786	42.76141632	38.42705427		ok	
7	DI Yogyakarta	153.9540009	62.26394606	20.56956733		ok	
8	Nusa Tenggara Barat	145.7744954	50.20002256	29.54853973		ok	
9	Gorontalo	98.55255248	33.48007137	86.66483081		ok	
10	Kalimantan Utara	83.27530907	35.97512713	86.89059903		ok	
11	Kalimantan Barat	97.46642448	32.63055294	77.31194086		ok	
12	Sulawesi Utara	81.79925061	28.48926965	88.70507934		ok	
13	Maluku	60.40976991	49.56064405	107.3927493		ok	
14	Papua Barat	82.11673398	41.42800106	92.32361399		ok	
15	Papua	41.48855987	80.65887456	132.2789949	ok	1	0

5. Penetuan pusat *cluster* baru Setelah didapatkan anggota dari setiap *cluster* kemudian pusat *cluster* baru dihitung berdasarkan data anggota tiap *cluster* yang sudah didapatkan menggunakan rumus yang sesuai dengan pusat anggota *cluster* sebagai berikut:

$$c_1 = \frac{(20.10), (34.43), (3.18), (11.64), (5.70), (30.42), (24.14), (6.80), (15.90), (5.80), (48.32), (34.48), (5.10), (15.30), (4.80)}{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1}$$

Hasil  $c_1$  =  
 $(20.10); (34.43); (3.18); (11.64); (5.70); (30.42); (24.14); (6.80); (15.90); (5.80); (48.32); (34.48); (5.10); (15.30); (4.80)$

$$c_2 = \frac{(23.50 + 12.40 + 32.20 + 12.50 + 31.20 + 22.90 + 19.70 + 16.70 + 22.90)}{9}; \dots; N$$

Hasil  $c_2$  =  
 $(21.56); (62.89); (3.90); (9.76); (4.05); (24.59); (80.04); (5.13); (11.84); (4.57); (50.22); (60.37); (5.11); (11.54); (3.64)$

$$c_3 = \frac{(45.30 + 41.00 + 39.60 + 55.40 + 38.30)}{5}, \frac{(94.12 + 100.00 + 100.00 + 100.00 + 100.00)}{5}, \dots; N$$

Hasil  $c_3$  =  
 $(43.92); (98.82); (2.48); (6.02); (2.09); (47.11); (100.00); (2.94); (7.96); (2.08); (69.02); (98.08); (3.64); (9.54); (3.48)$

Dari perhitungan di atas maka didapatkan pusat *cluster* baru dalam matrik tabel sebagai berikut:

**Tabel 3.3** Centroid baru iterasi ke-1

	C1	C2	C3
ASI 2016	20.10	21.56	43.92
IDL 2016	34.48	62.89	98.82
GB 2016	3.18	3.90	2.48
SP 2016	11.64	9.76	6.02
SK 2016	5.70	4.05	2.09
ASI 2017	30.42	24.59	47.11
IDL 2017	24.14	80.04	100.00
GB 2017	6.80	5.13	2.94
SP 2017	15.90	11.84	7.96
SK 2017	5.80	4.57	2.08
ASI 2018	48.32	50.22	69.02
IDL 2018	34.48	60.37	98.08
GB 2018	5.10	5.11	3.64
SP 2018	15.30	11.54	9.54
SK 2018	4.80	3.64	3.48

6. Hasil *iterasi* yang diperoleh sampai pada perhitungan *iterasi* ke-3 sebagai berikut:

**Tabel 3.4** Hasil Klasterisasi

Data ke	Provinsi	C1	C2	C3
1	Aceh	ok		
2	Sumatera Utara		ok	
3	Sumatera Selatan			ok
4	Bengkulu			ok
5	DKI Jakarta			ok
6	Jawa Barat			ok
7	DI Yogyakarta			ok
8	Nusa Tenggara Barat			ok
9	Gorontalo		ok	
10	Kalimantan Utara		ok	
11	Kalimantan Barat		ok	
12	Sulawesi Utara		ok	
13	Maluku		ok	
14	Papua Barat		ok	
15	Papua	ok		
		2	7	6

### Perhitungan nilai Sum of Squares Error (SSE)

Perhitungan nilai *Sum of Squares Error* (*SSE*) untuk menentukan jumlah *cluster* yang paling optimum atau yang terbaik. Langkah-langkah nya sebagai berikut:

1. Inisialisasi awal nilai *cluster* (2 *cluster*).
2. Menaikkan nilai *cluster* sampai jumlah *cluster* yang ditentukan (4 *cluster*).
3. Menghitung nilai SSE (*Sum of Squares Error*) dari setiap *cluster*.

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_K} \|X_i - C_K\|^2_2 \dots \dots \dots \dots \dots \dots \dots \quad (2)$$

**Tabel 3.5** *Sum of Squares Error* (SSE) pada 2 *cluster*

Data	M1	M2	MIN
1	1733.3402	9784.0183	1733.3402
2	577.6971	4971.7812	577.6971
3	5235.4833	178.2922	178.2922
4	4575.8029	707.8362	707.8362
5	6399.0632	52.0869	52.0869
...	...	...	...
	SSE	13142.4783	

Ambil nilai terkecil (MIN) dari M1 dan M2 masing-masing data, kemudian nilai terkecil di jumlah (SUM) dan hasil penjumlahan adalah nilai SSE (*Sum of Squares Error*).

4. Melihat hasil SSE dari setiap nilai *cluster* yang terkecil

**Tabel 3.6** *Sum of Squares Error* (SSE) pada 2 sampai 4 *cluster*

N cluster	SSE
2 cluster	13142.4783
3 cluster	8372.7104
4 cluster	6797.6534
Nilai SSE	6797.6534

5. Menetapkan nilai SSE terkecil sebagai *cluster* optimum.
6. Profiling *cluster* optimum.

## HASIL DAN PEMBAHASAN

### Pengumpulan Data

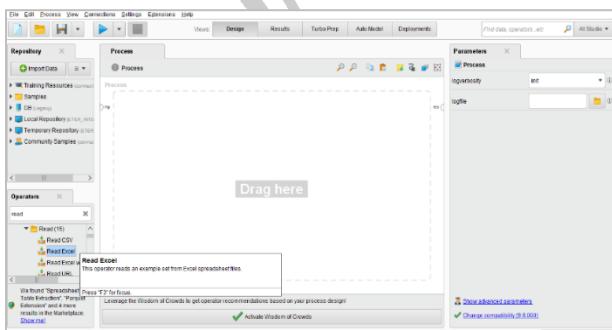
Sumber data utama yang digunakan dalam penelitian ini adalah data penyebab *stunting* yang diambil dari website Kementerian Kesehatan Republik Indonesia bagian publikasi data dan informasi. seluruh data terdiri dari 34 provinsi dengan jumlah keseluruhan data ada 510 record dari tahun 2016 sampai 2018.

**Tabel 4.1** Data set yang digunakan

Data ke	Provinsi	ASI 2016	IDL 2016	GB 2016	SP 2016	SK 2016	ASI 2017	IDL 2017	GB 2017	SP 2017	SK 2017	ASI 2018	IDL 2018	GB 2018	SP 2018	SK 2018
1	Aceh	23.50	69.57	2.55	7.56	3.51	22.99	52.17	5.90	12.20	3.70	60.84	21.74	6.70	16.00	5.00
2	Sumatera Utara	12.40	72.73	3.07	9.34	4.33	10.73	75.76	5.30	12.50	5.70	50.07	60.61	5.40	13.20	4.60
3	Sumatera Barat	37.60	73.48	2.13	6.66	1.88	36.02	84.21	3.30	9.30	2.80	77.09	47.37	3.50	9.60	2.90
4	Riau	25.00	66.67	1.96	7.32	3.24	28.57	75.00	4.20	11.20	4.00	78.33	33.33	4.30	10.30	4.20
5	Jambi	28.10	100.00	2.38	8.50	2.41	42.28	100.00	3.00	8.80	3.80	68.64	100.00	3.80	13.40	5.70
6	Sumatera Selatan	45.30	94.12	1.93	4.66	1.60	48.08	100.00	2.10	7.90	2.00	60.43	94.12	4.90	14.40	4.70
7	Bengkulu	32.20	80.00	1.30	6.61	1.30	25.55	100.00	2.30	8.60	2.70	76.38	100.00	2.80	9.80	3.50
8	Lampung	22.40	100.00	1.63	6.59	2.60	32.21	100.00	3.50	10.10	2.90	67.01	100.00	3.10	9.60	3.90
9	Kep. Bengkulu Bali	27.00	85.71	2.01	6.22	1.65	27.82	100.00	3.70	9.30	3.20	59.11	85.71	3.40	7.30	2.80
10	Kepulauan Riau	28.70	85.71	3.73	7.21	3.02	29.31	100.00	3.00	4.70	4.40	52.18	100.00	3.20	8.50	4.50
11	DKI Jakarta	41.00	100.00	2.99	6.28	3.50	46.60	100.00	3.00	7.20	2.60	71.88	100.00	2.30	6.10	3.90
12	Jawa Barat	39.60	100.00	2.35	6.13	1.64	38.23	100.00	2.90	8.40	1.60	57.97	96.30	2.60	11.70	3.20
13	Jawa Tengah	42.70	100.00	2.98	6.09	2.22	41.09	100.00	3.00	7.90	2.40	64.19	100.00	3.10	11.20	2.70
14	DI Yogyakarta	55.40	100.00	2.10	4.74	1.68	61.45	100.00	2.40	5.10	2.00	76.17	100.00	2.50	6.30	1.20
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
34	Papua	20.10	34.48	3.18	11.64	3.70	39.42	24.14	6.80	15.90	5.80	48.32	34.48	5.10	15.30	4.80

## Implementasi Algoritma K-Mans Menggunakan Rapidminer

Data diuji menggunakan tools rapidminer untuk mencari hasil dari 2 cluster sampai dengan 6 cluster, karena pengalokasian data pada 6 cluster sudah ada kelompok data yang hanya terdapat satu data (provinsi) dalam sebuah cluster. Data faktor penyebab *stunting* disimpan dengan format .xlsx (excel) agar bisa dianalisis menggunakan tools rapidminer. Tampilan awal tools rapidminer bisa dilihat pada gambar 4.1 sebagai berikut.



**Gambar 4.1.** Tampilan awal Rapidminer Studio

## Pengujian 2 Cluster Menggunakan Rapidminer

### Cluster Model

Cluster 0: 14 items  
Cluster 1: 20 items  
Total number of items: 34

**Gambar 4.2.** Hasil 2 cluster

Dengan menentukan 2 cluster pada data faktor penyebab *stunting* untuk mencari kelompok wilayah sejenis dapat terlihat jarak *centroid* antar cluster. Pada cluster 0 terdapat 14 items dan cluster 1 = 20 items. Jumlah data pada setiap clusternya terlihat pada tabel 4.2

**Tabel 4.2.** Centroid akhir 2 cluster

Attribute	cluster_0	cluster_1
ASI 2016	24.771	33.005
IDL 2016	62.439	92.988
GB 2016	3.917	3.027
SP 2016	9.624	7.463
SK 2016	4.088	2.284
ASI 2017	28.660	36.445
IDL 2017	71.540	99.375
GB 2017	5.329	3.735
SP 2017	12.536	9.450
SK 2017	4.414	2.755
ASI 2018	56.824	64.816
IDL 2018	46.366	93.523
GB 2018	5.379	3.785
SP 2018	12.321	10.250
SK 2018	3.786	3.470

## Pengujian 3 Cluster Menggunakan Rapidminer

### Cluster Model

Cluster 0: 18 items  
Cluster 1: 7 items  
Cluster 2: 9 items  
Total number of items: 34

**Gambar 4.3.** Hasil 3 cluster

Dengan menentukan 3 cluster maka jumlah data pada setiap clusternya akan terlihat, cluster 0 terdapat 18 items, cluster 1 = 7 items dan cluster 2 = 9 items yang terlihat pada pada gambar 4.3.

**Tabel 4.3. Centroid akhir 3 cluster**

Attribute	cluster_0	cluster_1	cluster_2
IDL 2015	95.753	65.390	61.402
GB 2015	2.972	3.367	4.258
SP 2015	7.229	8.721	10.311
SK 2015	2.172	3.650	4.251
ASI 2017	38.011	32.637	24.164
IDL 2017	99.306	61.914	85.351
GB 2017	3.444	5.229	5.633
SP 2017	8.878	12.814	12.778
SK 2017	2.561	3.871	4.856
ASI 2018	66.017	68.887	48.971
IDL 2018	94.538	32.503	65.599
GB 2018	3.628	5.543	5.211
SP 2018	10.167	13.486	11.122
SK 2018	3.461	4.100	3.989

**Pengujian 4 cluster menggunakan Rapidminer**

Dengan menentukan 4 *cluster* maka jumlah data pada setiap clusternya akan terlihat, *cluster 0* terdapat 2 *items*, *cluster 1* = 7 *items*, *cluster 2* = 18 *items* dan *cluster 3* = 7 *items* yang terlihat pada gambar 4.4

### Cluster Model

Cluster 0: 2 items  
 Cluster 1: 7 items  
 Cluster 2: 18 items  
 Cluster 3: 7 items  
 Total number of items: 34

**Gambar 4.4. Hasil 4 cluster**

Jarak antar clusternya juga terlihat seperti pada tabel 4.4 dimana *cluster* terpendek *cluster\_0* ada pada atribut Sangat Kurus 2018 (SK 2018) sedangkan *cluster* tertinggi *cluster\_0* pada atribut Air Susu Ibu 2018 (ASI 2018).

**Tabel 4.4. Centroid akhir 4 cluster**

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
ASI 2015	18.400	20.043	34.361	30.186
IDL 2015	35.420	63.547	95.753	70.669
GB 2015	4.590	3.664	2.972	3.866
SP 2015	11.985	9.790	7.229	8.764
SK 2015	7.365	3.557	2.172	3.454
ASI 2017	28.805	21.534	38.011	33.841
IDL 2017	43.890	87.380	99.306	71.731
GB 2017	6.300	5.486	3.444	5.186
SP 2017	13.100	13.100	8.878	12.400
SK 2017	5.850	4.729	2.561	3.714
ASI 2018	44.915	45.486	66.017	70.760
IDL 2018	44.515	70.426	94.538	33.700
GB 2018	6.250	4.900	3.628	5.657
SP 2018	13.900	10.886	10.167	12.929
SK 2018	4.400	3.343	3.461	3.986

**Pengujian 5 Cluster menggunakan Rapidminer**

Dengan menentukan 5 *cluster* maka jumlah data pada setiap clusternya akan terlihat, *cluster 0* terdapat 2 *items*, *cluster 1* = 5 *items*, *cluster 2* = 16 *items*, *cluster 3* = 7 *items* dan *cluster 4* = 4 *items* yang terlihat pada gambar 4.5

### Cluster Model

Cluster 0: 2 items  
 Cluster 1: 5 items  
 Cluster 2: 16 items  
 Cluster 3: 7 items  
 Cluster 4: 4 items  
 Total number of items: 34

**Gambar 4.5. Hasil 5 cluster**

Jarak antar clusternya juga terlihat seperti pada tabel 4.5 dimana *cluster* terpendek *cluster\_0* ada pada atribut Sangat Kurus 2018 (SK 2018) sedangkan *cluster* tertinggi *cluster\_0* pada atribut Air Susu Ibu 2018 (ASI 2018).

**Tabel 4.5. Centroid akhir 5 cluster**

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
ASI 2015	18.400	21.300	35.381	30.186	21.550
IDL 2015	35.420	74.658	97.009	70.669	60.490
GB 2015	4.590	3.846	2.865	3.866	3.520
SP 2015	11.985	9.816	7.044	8.764	9.467
SK 2015	7.365	3.076	2.161	3.454	3.513
ASI 2017	28.805	22.912	39.190	33.941	23.335
IDL 2017	43.890	100	99.219	71.731	77.915
GB 2017	6.300	5.680	3.269	5.186	4.925
SP 2017	13.100	13.020	8.444	12.400	12.825
SK 2017	5.850	4.020	2.463	3.714	4.925
ASI 2018	44.915	54.538	66.809	70.760	40.870
IDL 2018	44.515	79.946	96.088	33.700	64.383
GB 2018	6.250	5.220	3.525	5.557	4.275
SP 2018	13.900	10.940	10.187	12.929	10.375
SK 2018	4.400	3.540	3.469	3.986	3.125

**Pengujian 6 Cluster menggunakan Rapidminer**

Pada pengujian di *cluster* = 6 terdapat *cluster* yang dalamnya hanya berisi 1 *item*, yang berarti *cluster* tersebut hanya berisi 1 provinsi saja, terlihat pada gambar 4.6, maka pengujian dihentikan pada penentuan *cluster* = 6, karena apabila pengujian dilanjutkan ke *cluster* = 7 akan ada

*cluster* yang hanya memiliki 1 *items* (provinsi), begitu juga seterusnya.

### Cluster Model

```

Cluster 0: 18 items
Cluster 1: 2 items
Cluster 2: 6 items
Cluster 3: 3 items
Cluster 4: 1 items
Cluster 5: 4 items
Total number of items: 34

```

Gambar 4.6. Hasil 6 cluster

Adapun jarak antar clusternya juga terlihat seperti pada tabel 4.6 dimana *cluster* terpendek *cluster\_0* ada pada atribut Sangat Kurus 2016 (SK 2016) sedangkan *cluster* tertinggi *cluster\_0* pada atribut Imunisasi Dasar Lengkap 2017 (IDL 2017).

Tabel 4.6. Centroid akhir 6 cluster

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5
ASI 2016	34.361	23.950	18.183	33.557	20.100	27.650
IDL 2016	95.753	38.180	67.472	76.147	34.480	66.560
GB 2016	2.972	5.035	3.597	4.613	3.180	3.305
SP 2016	7.229	11.310	9.707	11.097	11.640	7.015
SK 2016	2.172	5.785	3.727	3.137	5.700	3.692
ASI 2017	38.011	32.485	18.827	39.027	30.420	30.128
IDL 2017	99.306	71.820	88.610	88.800	24.140	60.430
GB 2017	3.444	6.150	5.850	4.900	6.800	5.400
SP 2017	8.878	10.800	13.400	12.400	15.900	12.400
SK 2017	2.561	4.600	4.967	3.200	5.800	4.100
ASI 2018	66.017	47.900	44.018	73.400	48.320	68.780
IDL 2018	94.538	57.275	72.163	41.187	34.480	28.085
GB 2018	3.628	4.900	5.317	5	5.100	5.975
SP 2018	10.167	9.850	11.557	12.400	15.300	13.325
SK 2018	3.461	2.550	3.717	3.367	4.800	4.450

## Menghitung Nilai Sum of Squares Error (SSE) Setiap Jumlah Cluster

Selanjutnya akan di cari nilai *Sum of Squares Error (SSE)* dari masing-masing *cluster* yaitu 2 *cluster* sampai dengan 6 *cluster*. Untuk menghitung nilai SSE dari sebuah *cluster* dengan menggunakan rumus (2) seperti yang ada di bab sebelumnya.

### Nilai SSE 2 Cluster

Tabel 4.7. Hasil perhitungan SSE 2 cluster

Data	M1	M2	MIN	Data	M1	M2	MIN
1	1106.559678	8275.5602	1106.559678	18	5937.421563	343.6911108	343.6911108
2	850.6913061	3395.3303	850.6913061	19	1584.295835	8587.0963	1584.295835
3	963.0925633	2907.5225	963.0925633	20	721.4323061	3254.8174	721.4323061
4	680.8813776	5226.4622	680.8813776	21	2409.436792	494.6944607	494.6944607
5	5467.049006	182.6950	182.6949608	22	3278.600635	201.4756307	201.4756307
6	4983.505578	342.9558308	342.9558308	23	4378.306863	150.1634	150.1633998
7	4506.584335	472.7366	472.7366107	24	944.7805776	4455.6702	944.7805776
8	5260.141649	231.1082	231.1082008	25	836.0295918	3432.0955	836.0295918
9	2976.17332	269.9222	269.922207	26	3041.973292	1985.8131	1985.813061
10	4359.16452	355.7880	355.7880407	27	4583.484349	243.5761408	243.5761408
11	6006.988706	336.5657	336.5657108	28	2222.559692	984.8894	984.8894107
12	5086.949478	160.1784	160.1783608	29	1752.17662	2767.9412	1752.17662
13	5703.5102	222.4602	222.4602308	30	1753.79106	4420.9498	1753.79106
14	7638.762892	1399.5216	1399.52161641	31	1159.026006	7008.6582	1159.026006
15	5358.158035	229.7398	229.7395808	32	794.9082633	5978.0970	794.9082633
16	3642.368092	331.1610	331.1610108	33	2378.041063	4115.4528	2378.041063
17	5374.655449	186.1415	186.1414608	34	3299.60192	13162.9970	3299.60192

Tabel 4.8. Nilai SSE 2 cluster

SSE	27950.5852

### Nilai SSE 3 Cluster

Tabel 4.9. Hasil perhitungan SSE 3 cluster

Data	M1	M2	M3	MIN	Data	M1	M2	M3	MIN
1	8643.601697	433.7822	338.4217	433.7822204	18	247.4039639	7476.2684	4541.6262	247.4039639
2	3765.637353	2169.1078	501.7245	501.7244802	19	8926.147853	599.2667	4080.0669	599.2666776
3	3079.048331	976.8122	1893.0280	976.8122061	20	3523.409775	1259.4418	1180.3326	1180.33258
4	5522.652397	325.4839	2226.6128	325.4839204	21	649.558775	4136.2507	1182.6948	649.558775
5	131.9068528	7329.3187	3797.4548	131.9068528	22	272.3995417	5096.4543	1890.6235	272.3995417
6	287.752417	6681.7748	3593.3653	287.7524517	23	145.6797306	602.3300	3020.5723	145.6797306
7	551.9846083	6398.5049	2831.7464	551.9846083	24	4858.260319	2034.4782	923.817958	4858.260319
8	230.8503972	7302.8505	3405.8831	230.8503972	25	3840.760219	2517.1081	139.6605691	3840.760219
9	396.6207417	4888.8922	1485.2793	396.6207417	26	2325.802708	5524.7638	1047.794869	2325.802708
10	456.7730972	6808.4296	2214.1337	456.7730972	27	181.3129083	5885.2452	3582.6701	181.3129083
11	224.7343861	7464.4874	4548.5097	224.7343861	28	1213.665319	4150.3474	842.1421	1213.665319
12	120.6747972	7015.4235	3422.7568	120.6747972	29	3163.890886	3805.5092	440.5288	3163.890886
13	140.6736528	7530.4095	4079.7129	140.6736528	30	4548.979742	1157.7114	3405.6119	4548.979742
14	1187.654808	8909.2366	6540.2532	1187.654808	31	7549.178197	2335.3919	1314.5155	7549.178197
15	176.0693194	6976.6632	3918.5751	176.0693194	32	6253.015864	408.9107	2548.6730	6253.015864
16	320.7020417	5161.6838	2597.3347	320.7020417	33	4483.342397	4567.9503	1176.6530	4483.342397
17	162.4071861	7399.8625	3545.5223	162.4071861	34	13718.79578	2932.22629	5517.1686	13718.79578

Jumlah MIN

20286.52318

### Nilai SSE 3 cluster

SSE	20286.52318

### Nilai SSE 4 cluster

Tabel 4.11. Hasil perhitungan SSE 4 cluster

Data	M1	M2	M3	M4	MIN
1	2116.079625	3936.541102	8643.6017	805.440402	805.440402
2	3077.214325	521.5365163	3765.6374	2033.974445	521.5365163
3	4679.145325	2202.919373	3079.0483	491.4242735	491.4242735
4	3303.616025	2708.268002	5522.6524	156.6504163	156.6504163
5	1131.9068528	3437.38493	131.9069	6160.872207	131.9068528
6	10546.16043	3314.608316	287.7525	5594.555845	287.752417
7	9557.568925	2483.902545	551.9846	5445.222131	551.9846083
8	1110.050453	2982.320088	230.8505	6169.187245	230.8503972
9	7805.890325	1212.085331	396.6207	3976.734316	396.6207417
10	9079.570925	1799.032916	456.7731	5882.513816	456.7730972
11	1213.507273	4215.134847	224.7344	6430.822773	224.7343861
12	10852.68763	3037.975688	120.6748	5897.904159	120.6747972
13	11664.47573	3698.604359	140.6737	6364.979888	140.6736528
14	14076.16913	6291.451259	1187.6548	7624.37861	1187.654808
15	11458.05773	3568.533731	176.0693	5793.103631	176.0693194
16	8569.310725	2359.019702	320.7020	4314.256859	320.7020417
17	11239.67893	3107.179273	162.4072	6260.412002	162.4071861
18	12107.71033	4201.780359	247.4040	6256.912888	247.40439639
19	2157.155725	4851.035599	8926.1479	998.2381735	998.2381735
20	4176.596925	150.1671959	3523.4098	702.5638306	702.5638306
21	7110.29025	997.5130163	649.5588	3248.370688	649.5588775
22	8368.084725	1644.1613616	272.3995	4112.795245	272.3995417
23	10224.85583	2717.251273	145.6797	4887.811259	145.6797306
24	2012.504025	1219.223473	4858.2603	2048.556431	2048.556431
25	2841.952852	195.5484592	3840.7602	2309.673645	195.5484592
26	5923.860425	825.7079592	2325.8027	4974.226088	825.7079592
27	10638.41943	3358.077073	181.3129	4714.444802	181.3129083
28	6570.736125	624.5570163	1213.6653	3351.728202	624.5570163
29	4931.586625	335.4858878	3163.8909	3247.929645	335.4858878
30	5914.187925	3866.582473	4548.9797	600.541802	600.541802
31	525.153925	1674.192845	7549.1782	2821.783345	525.153925
32	2899.443725	2999.838916	6253.0159	478.0615878	478.0615878
33	3523.757725	1052.317573	4483.3424	4642.221331	1052.317573
34	525.153925	6263.896488	13718.7958	4232.325873	525.153925

Jumlah MIN

15942.76457

### Tabel 4.12. Nilai SSE 4 cluster

SSE	15942.76457


<tbl\_r cells="2" ix="2" maxc

## Nilai SSE 5 Cluster

Tabel 4.13. Hasil perhitungan SSE 5 cluster

Data	M1	M2	M3	M4	M5	MIN
1	2116.079625	5785.015476	9005.5081	805.440402	3014.814869	805.440402
2	3077.214325	1225.512696	6075.0700	4340.647424	3668.498188	1225.512696
3	4679.145325	2297.277736	2272.9451	1490.767376	3345.946238	1490.767376
4	3303.616025	3492.907336	2763.7177	1017.60062	2413.894338	1017.60062
5	11319.14133	1696.119876	10094.8984	11763.1822	13262.40566	1696.119876
6	10546.16043	1903.013176	9116.8186	10802.55384	12125.43121	1903.013176
7	9557.568925	1089.978216	10580.8571	11049.53213	12989.00372	1089.978216
8	11028.05463	1312.743376	10232.4260	11773.49724	12840.59856	1312.743376
9	7805.890325	280.949236	7746.7539	8617.898316	8896.039063	280.949236
10	9079.570925	714.495616	10510.8344	11486.82382	11420.47656	714.495616
11	12135.02723	2371.433596	10143.1758	12035.13277	14005.94011	2371.433596
12	10852.687763	1539.822336	9837.3383	11252.83416	12217.86866	1539.822336
13	11664.47573	2000.772836	10080.8998	11969.2898	13434.16886	2000.772836
14	14076.16913	4290.955456	11031.1728	13228.68862	16019.24551	4290.955456
15	11458.05773	1736.257716	9636.3000	11220.15163	13114.52251	1736.257716
16	8569.310725	1382.938436	7931.2525	9076.066859	10052.53016	1382.938436
17	11239.67893	1454.164676	10147.3275	11864.722	12914.52481	1454.164676
18	12170.71033	2253.626756	10177.2752	11861.22289	14093.44391	2253.626756
19	2157.155723	6639.648436	5233.53891	1700.546173	3304.012313	1700.546173
20	4176.596925	1802.749756	2788.1493	2455.637831	2948.944363	1802.749756
21	7110.290525	210.006036	6659.1707	7408.298688	7689.844313	210.006036
22	8368.084725	616.956136	7382.9501	8680.493245	9150.062063	616.956136
23	10224.85583	1165.918396	8238.4173	9818.121259	11178.12616	1165.918396
24	2012.504025	2368.758556	7438.9191	4956.866431	4533.527263	2012.504025
25	2841.952825	1075.388056	6458.0836	5217.983645	3703.406063	1075.388056
26	5923.860425	599.661676	11040.8000	10060.23009	9249.924563	599.661676
27	10638.41943	1736.470756	7753.9338	9476.254802	11440.98906	1736.470756
28	6570.736125	85.506236	6872.6856	7370.116202	6992.518963	85.506236
29	4931.586625	498.009336	7054.5184	6605.979645	5224.493663	498.009336
30	5914.187925	3952.426596	1910.5076	1711.293802	3973.184263	1711.293802
31	525.153925	3724.092916	9185.1756	5362.763345	3868.454013	525.153925
32	2899.443725	4237.894136	3083.1532	1364.371588	2088.803613	1364.371588
33	3523.757725	2171.443696	10308.0647	8690.959331	6509.570363	2171.443696
34	525.153925	9593.507536	11534.8458	5420.587873	4916.042613	525.153925
Jumlah MIN						
46367.52392						

Tabel 4.14. Nilai SSE 5 cluster

SSE	46367.52392

## Nilai SSE 6 Cluster

Tabel 4.15. Hasil perhitungan SSE 6 cluster

Data	M1	M2	M3	M4	M5	M6	MIN
1	8643.601697	2970.3028	4235.4305	2173.640856	2446.6513	258.0908375	258.0908375
2	3765.637353	1862.7764	467.5788	2319.472202	5286.0514	2298.659038	467.5788167
3	3079.48331	2619.054	2488.6710	140.0180222	6620.4852	1233.779738	140.0180222
4	5522.652397	2374.3504	3013.9528	527.2874889	4636.5832	357.4733875	527.2874889
5	131.9068528	7045.3558	3261.1387	4302.428856	15054.3261	8033.504588	131.9068528
6	287.7525417	6260.0127	3275.9662	3773.38422	14130.6494	7439.235938	287.7525417
7	351.9846083	3373.7805	2420.1738	3908.014536	13245.7462	707.923588	351.9846083
8	230.8509725	6864.1043	2713.6646	4462.795956	1482.04546	7927.781488	230.8509725
9	396.6207417	4084.7958	1097.6920	2693.649656	517.848538	399.6207417	517.848538
10	456.7730972	5005.6462	1628.6447	4393.666487	12981.1044	7478.400088	456.7730972
11	224.7348611	7583.6982	4133.1788	4421.391856	13821.4221	8416.696738	224.7348611
12	120.6747972	6599.066	2870.4781	4116.689956	14562.0936	7712.615388	120.6747972
13	140.6736258	7214.3492	3568.2576	4437.239458	15392.8311	8289.585988	140.6736258
14	1187.654808	9202.7647	6359.6183	5347.383289	17641.5156	9018.923588	1187.654808
15	176.0693194	7192.5901	3402.6864	3978.578189	15106.4459	7632.048488	176.0693194
16	320.7020417	5284.9693	2244.0631	3046.016322	11629.3535	5744.238038	320.7020417
17	162.4071861	6926.419	2888.0184	4515.39558	15060.1793	8047.97658	162.4071861
18	247.4039639	7705.3868	4090.8432	4283.026722	15059.9411	8216.123838	247.4039639
19	8926.147853	2491.2044	5477.5995	2043.61598	2435.4475	693.0055875	693.0055875
20	3523.409775	2047.0726	1713.8710	300.5233889	6499.4193	1482.894938	300.5233889
21	649.558775	3678.1483	895.2386	2035.196856	10487.3395	4623.551838	649.558775
22	272.3995417	4624.3846	1523.0345	2665.341556	11869.5004	5677.182688	272.3995417
23	145.6797306	6232.5897	2357.0944	3261.012789	13794.6167	6586.710888	145.6797306
24	4858.260319	230.4212	1659.4986	2125.050122	4153.6311	2469.969698	230.4212
25	3840.760219	950.6963	286.4371	2300.513022	5444.6224	2795.344888	286.4371167
26	2325.802708	2417.3013	869.8239	4033.874722	9586.0161	6158.290388	869.8231167
27	181.3129083	6526.2607	3280.7353	2933.408122	14058.2287	6529.023088	181.3129083
28	1213.665319	3338.3116	478.8206	2386.124989	9975.7185	4554.731388	478.8204167
29	3163.890886	2335.5982	252.0337	2761.162889	8240.0688	4091.805488	252.0336833
30	4548.979742	4224.8573	4162.4948	259.5867222	7481.6629	1355.058888	259.5867222
31	7549.178197	230.4212	2036.6821	3796.362789	2100.6157	2569.649538	230.4212
32	6253.015864	2993.3198	3198.6768	1341.852156	3738.9524	309.0194375	309.0194375
33	4483.342397	2253.2806	997.2544	4880.619022	5935.544	4942.223888	997.25435
34	13718.79578	2896.7022	6852.6786	6633.813089	0	2910.011238	0
Jumlah MIN							
12015.66863							

Tabel 4.16. Nilai SSE 6 cluster

SSE	12015.66863

## Penentuan Jumlah Cluster Optimum

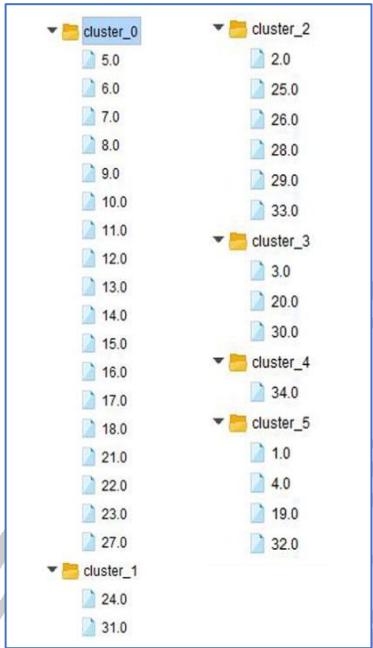
Nilai SSE kemudian dapat digunakan untuk validasi metode *clustering K-Means* karena meminimalkan SSE (*chohesien*) *ekuivalen* dengan memaksimalkan SSB (*separation*) (Nurwidodo dan Hariadi, 2013). Jumlah cluster yang memiliki nilai *Sum of Squares Error* (SSE) terkecil yang diambil sebagai cluster optimum, karena semakin kecil nilai SSE maka data dalam clusternya akan semakin mirip. Pada tabel 4.22 menunjukkan 6 cluster adalah jumlah cluster optimum dengan nilai SSE terendah yaitu 12015.66863.

Tabel 4.17. Nilai SSE tiap jumlah cluster

CLUSTER	SSE
2 Cluster	27950.5852
3 Cluster	20286.52318
4 Cluster	15942.76457
5 Cluster	46367.52392
6 Cluster	12015.66863

## Profiling Cluster Optimum

Dari hasil pengujian *clustering* menggunakan algoritma *K-Means* dan mencari cluster optimum menggunakan nilai SSE, diketahui cluster optimum berada di 6 cluster. Untuk mengetahui karakteristik dari masing-masing cluster, dilakukan *profiling cluster* data faktor penyebab *stunting* setiap provinsi yang masuk ke dalam kelompok cluster 1 s/d cluster 6 berdasarkan perhitungan yang telah telah dilakukan sebelumnya. Adapun hasil 6 cluster adalah sebagai berikut:



**Gambar 4.7.** Hasil 6 cluster pada Rapidminer

**Tabel 4.18.** Karakteristik cluster 1

Cluster 1	2016	2017	2018	Rata-rata	Selisih
ASI Terendah	22.40	25.55	52.18	33.38	31.78
ASI Tertinggi	55.40	61.45	78.63	65.16	
IDL Terendah	80.00	87.50	78.57	82.02	17.98
IDL Tertinggi	100.00	100.00	100.00	100.00	
GB Terendah	0.98	2.00	2.00	1.66	4.19
GB Tertinggi	5.65	6.00	5.90	5.85	
SP Terendah	4.66	4.70	5.60	4.99	8.68
SP Tertinggi	11.20	15.40	14.40	13.67	
SK Terendah	1.16	1.60	1.20	1.32	3.21
SK Tertinggi	3.50	4.40	5.70	4.53	

Tabel 4.23 dan tabel 4.24 menunjukkan pada cluster 1 terdapat 18 anggota provinsi dengan jumlah rata-rata persentase pemberian Air Susu Ibu (ASI) terendah selama tahun 2016 sampai 2018 sebesar 33.38% dan rata-rata ASI tertinggi sebesar 65.16%, adapun selisih dari persentase tertinggi ke terendah sebesar 31.78%. Selanjutnya untuk jumlah rata-rata Imunisasi Dasar Lengkap (IDL) terendah adalah 82.02% dan jumlah rata-rata pada IDL tertinggi mencapai 100% dengan selisih

17.98%. Sedangkan untuk jumlah rata-rata pada atribut Gizi Buruk (GB) terendah dari tahun 2016 sampai 2018 adalah 1.66% dan jumlah rata-rata GB tertinggi adalah 5.85% dengan selisih sebesar 4.19%. Untuk rata-rata atribut Sangat Pendek (SP) terendah adalah 4.99% dan pada SP tertinggi adalah 13.67% dengan selisih sebesar 8.68%. Kemudian pada atribut Sangat Kurus (SK) terendah memiliki jumlah rata-rata sebesar 1.32% dan 4.53% untuk SK tertingginya dengan jarak selisih 3.21%.

**Tabel 4.19.** Karakteristik cluster 6

Cluster 6	2016	2017	2018	Rata-rata	Selisih
ASI Terendah	23.50	22.99	60.84	35.78	17.48
ASI Tertinggi	38.30	43.14	78.33	53.26	
IDL Terendah	50.00	52.17	21.74	41.30	21.47
IDL Tertinggi	80.00	75.00	33.33	62.78	
GB Terendah	1.80	4.10	4.30	3.40	3.80
GB Tertinggi	6.91	7.40	7.30	7.20	
SP Terendah	4.87	8.20	10.30	7.79	6.31
SP Tertinggi	8.31	18.00	16.00	14.10	
SK Terendah	2.19	2.70	4.00	2.96	2.65
SK Tertinggi	5.83	6.00	5.00	5.61	

Tabel 4.33 dan tabel 4.34 menunjukkan pada cluster 6 terdapat 4 anggota provinsi dengan jumlah rata-rata persentase pemberian Air Susu Ibu (ASI) terendah selama tahun 2016 sampai 2018 sebesar 35.78% dan rata-rata ASI tertinggi sebesar 53.26%, adapun selisih dari persentase tertinggi ke terendah sebesar 17.48%. Selanjutnya untuk jumlah rata-rata Imunisasi Dasar Lengkap (IDL) terendah adalah 41.30% dan jumlah rata-rata pada IDL tertinggi mencapai 62.78% dengan selisih 21.47%. Sedangkan untuk jumlah rata-rata pada atribut Gizi Buruk (GB) terendah dari tahun 2016 sampai 2018 adalah 3.40% dan jumlah rata-rata GB tertinggi adalah 7.20% dengan selisih sebesar 3.80%. Untuk rata-rata atribut Sangat Pendek (SP) terendah adalah 7.79% dan pada SP tertinggi adalah 14.10% dengan selisih sebesar 6.31%. Kemudian

pada atribut Sangat Kurus (SK) terendah memiliki jumlah rata-rata sebesar 2.96 % dan 5.61% untuk SK tertingginya dengan jarak selisih 2.65%.

Jika dirata-ratakan, maka akan didapat provinsi yang berada dalam *cluster* 1 memiliki rata-rata persentase yang paling tinggi pada atribut ASI dan IDL serta persentase paling rendah untuk atribut GB, SK dan SP, sehingga resiko terjadinya stunting sangat rendah daripada provinsi yang berada di *cluster* lainnya. Sedangkan provinsi yang berada di *cluster* 5 memiliki rata-rata persentase yang paling rendah pada atribut ASI dan IDL serta persentase paling tinggi untuk atribut GB, SK dan SP, sehingga sangat rentan terjadi stunting daripada provinsi yang berada di *cluster* lainnya. Seperti yang terlihat pada tabel dibawah ini.

**Tabel 4.20.** Rata-rata dan selisih nilai pada atribut

Cluster	Rata-rata ASI dan IDL	Selisih dengan nilai (ASI, IDL) tertinggi	Rata-rata GB, SK dan SP	Selisih dengan nilai (GB, SK, SP) terendah	Rata-rata (ASI, IDL) dan (GB, SK, SP)
C1	71.33	0.00	4.95	0.00	76.28
C2	45.27	26.06	6.64	1.70	51.91
C3	51.55	19.79	6.85	1.90	58.40
C4	58.35	12.98	6.68	1.73	65.03
C5	31.99	39.34	8.25	3.30	40.24
C6	46.94	24.39	6.63	1.68	53.57

## KESIMPULAN

Berdasarkan uraian permasalahan dan pembahasan di atas, maka diperoleh kesimpulan sebagai berikut: (1) Penerapan algoritma *K-Means* pada data faktor penyebab *stunting* untuk mengelompokkan provinsi di Indonesia menghasilkan 6 *cluster* optimum dengan nilai *Sum of Squares Error* (SSE) adalah 12015.66863, (2) Hasil pengelompokan pada *cluster* 1 terdapat 18 provinsi yaitu Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa tengah, Jawa Timur, Banten, Bali, NTB, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur dan Sulawesi Selatan. Pada *cluster* 2 terdapat 2 provinsi

yaitu Kalimantan Utara dan Maluku. Pada *cluster* 3 terdapat 6 provinsi yaitu Sumatera Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo dan Papua Barat. Pada *cluster* 4 terdapat 3 provinsi yaitu Sumatera Barat, Kalimantan Barat dan Sulawesi Barat. Pada *cluster* 5 terdapat 1 provinsi yaitu Papua. Pada *cluster* 6 terdapat 4 provinsi yaitu Aceh, Riau, NTT dan Maluku Utara, (3) Provinsi yang berada di *cluster* 1 memiliki rata-rata persentase yang paling tinggi pada atribut ASI dan IDL serta persentase paling rendah untuk atribut GB, SK dan SP, sehingga resiko terjadinya stunting sangat rendah daripada provinsi yang berada di *cluster* lainnya, (4) Provinsi yang berada di *cluster* 5 memiliki rata-rata persentase yang paling rendah pada atribut ASI dan IDL serta persentase paling tinggi untuk atribut GB, SK dan SP, sehingga sangat rentan terjadi stunting daripada provinsi yang berada di *cluster* lainnya.

## DAFTAR PUSTAKA

- Buletin Jendela (*Situasi Balita Pendek (Stunting) Di Indonesia*). 2018. Jakarta. Pusdatin Kementrian Kesehatan RI.
- Dhuhiba, Whinda Mega P. 2015. *Clustering Dengan Metode K-Means Untuk Mengetahui Status Gizi Balita*. *Jurnal Informatika*. Vol.15, hal.160-174.
- Fitri. 2013. *Berat Lahir Sebagai Faktor Dominan Terjadinya Stunting Pada Balita (12–59 Bulan) di Sumatera (Analisis Data Riskesdas 2010)*. *Jurnal Photon*. Vol.4, hal.77-88.
- Irfani, Erni & Rani, Siti Sulistia. 2018. Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita. *Jurnal Sistem dan Teknologi Informasi*. Vol.6, hal.165-172.
- Izza, Nailul. 2019. *Analisis Pola Terjadinya Stunting Balita Di Indonesia Dengan Menggunakan Metode DBSCAN Clustering*. Tesis. Universitas Airlangga.
- Kaur, K., Dhaliwal, D.S. & Vohra, K.R. 2013. *Statistically Refining the Initial*

- Points for K-Means Clustering Algorithm. International Journal of Advanced Research in Computer Engineering & Technology.* Vol.2, pp.2972-2977.
- Kementrian Kesehatan Fokus pada Pencegahan Stunting, [online]. <https://indonesia.go.id/narasi/indonesia-dalam-angka/sosial/kementerian-kesehatan-fokus-pada-pencegahan-stunting>, [diakses 01 Agustus 2020].
- Mahrus, M. A & Khotimah, B. K. 2012. *Pengelompokan Tingkat Kesehatan Masyarakat Menggunakan K-Means Berbasis OWA (Ordered Weighted Averaging)* Vol .1: 47-56.
- Nainggolan, Rena & Lumbantoruan, Gortap. 2018. *Optimasi Performa Cluster K-Means Menggunakan Sum of Squares Error (SSE)*. *Jurnal Methomika*. Vol.2, Hal. 103-108.
- Ni'mah, Khoirun & Nadhiroh, Siti Rahayu. 2015. *Faktor Yang Berhubungan Dengan Kejadian Stunting Pada Balita*.
- Nurwidodo & Hariadi, M. 2013. *Sistem Pendukung Keputusan Multidimensi Menggunakan K-Means Clustering Berbasis Mahalanobis Distance*. *Prosiding Seminar Nasional Manajemen Teknologi XVII*. pp 20-4-20-11.
- Prayogo, Ari. dkk. 2009. *Kelengkapan Imunisasi dasar Pada Anak Usia 1-5 Tahun*. *Sari Pedriati*. Vol. 11, hal.15-19.
- Profil Kesehatan Indonesia Tahun 2016*. 2017. Jakarta. Kementrian Kesehatan RI.
- Profil Kesehatan Indonesia Tahun 2016*. 2018. Jakarta. Kementrian Kesehatan RI.
- Profil Kesehatan Indonesia Tahun 2016*. 2019. Jakarta. Kementrian Kesehatan RI.
- Profil Kesehatan Indonesia Tahun 2016*. 2020. Jakarta. Kementrian Kesehatan RI.
- Rahman, Aulia T. dkk. 2017. *Coal Trade Data Clustering Using K-Means (Case Study PT. Global Bangkit Utama)*. *Jurnal Ilmiah Teknologi dan Informasi*. Vol.6, hal.24-31.
- Saputri, Rini Archda & Tumanger, Jeki. 2019. *Hulu Hilir Penanggulangan Stunting di Indonesia*. *Jurnal Of Political Issues*. Vol.1, hal.1-9.
- Suryawati, Ida. Dkk. 2016. *Cakupan Imunisasi Dasar Anak Ditinjau Dari Pendekatan Health Belief Model Immunization Covarge In Children Based on Health Believ Model*. *Jurnal Ilmu Keperawatan*. Vol.4, hal.115-125.
- Swatma, D. dkk. 2016. *Analisis Faktor Risiko BBLR, Panjang Badan Bayi Saat Lahir dan Riwayat Imunisasi Dasar Terhadap Kejadian Stunting Pada Balita Usia 12-36 Bulan di Wilayah Kerja Puskesmas Kandai Kota Kendari*.
- Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K). 2017. *100 Kabupaten/Kota Prioritas Untuk Intervensi Anak Kerdil (Stunting)*. Vol.2. Sekretariat Wakil Presiden Republik Indonesia. Jakarta Pusat.