

Klasterisasi Angka IPM (Indeks Pembangunan Manusia) Tahun 2010 – 2022 pada Kabupaten/Kota di Indonesia menggunakan Metode Klasifikasi Hierarki

Haryoso Wicaksono

STMIK IM, Bandung, Indonesia

yosnex@yahoo.com

Informasi Artikel

Kata Kunci:

Cluster, Klasifikasi, IPM

Histori Artikel:

Disubmit 16 Maret 2023

Direvisi 28 April 2023

Diterima 5 Juni 2023

Tersedia daring 31 Juli 2023

Sitasi:

H. Wicaksono, "Klasterisasi Angka IPM (Indeks Pembangunan Manusia) Tahun 2010 – 2022 pada Kabupaten/Kota di Indonesia menggunakan Metode Klasifikasi Hierarki," Pros. Semnastek Univ. Suryakancana, 2023.

* Penulis Korespondensi.

Haryoso Wicaksono

yosnex@yahoo.com

Abstrak

Angka IPM (Indeks Pembangunan Manusia) bisa di pakai untuk mengidentifikasi kualitas hidup masyarakat pada suatu wilayah Kabupaten/ Kota di seluruh Indonesia. Berdasarkan data sekunder angka IPM dari BPS, bisa di ketahui Kabupaten/Kota mana yg mempunyai tingkat kualitas hidup yang tinggi atau rendah. Salah satu caranya adalah dengan Klasterisasi atau Cluster atau Klasifikasi. Metode Klasifikasi yang di pilih menggunakan metode Clustering Hierarki. Data yang diolah sebanyak 515 record. Jumlah Cluster yang direncanakan adalah 2, 3, 4, 5, 6 atau 7 mengacu kepada Angka IPM Kabupaten/Kota seluruh Indonesia periode 2010-2022. Di awali dengan persiapan data terkait kelengkapan data, maka di lakukan uji MSA (Missing Value Analysis). Dengan menggunakan aplikasi SPSS, proses Klasterisasi menggunakan data angka IPM pasca analisa MVA. Hasil proses klasterisasi menghasilkan sebaran tiap wilayah berdasarkan kelompok klasternya. Misal, pada 5 cluster terdapat sebaran 5 level angka IPM pada wilayah Kabupaten/Kota , yaitu Sangat Tinggi (432 Kabupaten/Kota), Tinggi (60 Kab/Kota), Sedang (9 Kab/Kota), Rendah (13 Kab/Kota) & Sangat Rendah (1 Kab/Kota). Jumlah Cluster yang di hasilkan sudah sesuai rencana, yaitu jumlah cluster sebanyak 2, 3, 4, 5, 6 atau 7. Selain proses klaster, masih di perlukan analisa data sortasi & pivot table menggunakan MS Excel untuk meningkatkan analisis sesuai kepentingan tertentu. Misal pada cluster 5, pada level Sangat Tinggi, wilayah Kabupaten/Kota mana yg mempunyai data angka IPM tertinggi.

1. Pendahuluan

Mengutip isi *Human Development Report (HDR)* pertama tahun 1990, Pembangunan Manusia adalah suatu proses untuk memperbanyak pilihan-pilihan yang dimiliki oleh manusia. Diantara banyak pilihan tersebut, pilihan yang terpenting adalah untuk berumur panjang dan sehat, untuk berilmu pengetahuan, dan untuk mempunyai akses terhadap sumber daya yang dibutuhkan agar dapat hidup secara layak. Indeks Pembangunan Manusia (IPM) mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup umur panjang dan sehat; pengetahuan, dan kehidupan yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Untuk mengukur dimensi kesehatan, digunakan angka harapan hidup waktu lahir. Selanjutnya untuk mengukur dimensi pengetahuan digunakan gabungan indikator angka melek huruf dan rata-rata lama sekolah. Adapun untuk mengukur dimensi hidup layak digunakan indikator kemampuan daya beli masyarakat terhadap sejumlah kebutuhan pokok yang dilihat dari rata-rata besarnya pengeluaran per kapita sebagai pendekatan pendapatan yang mewakili capaian pembangunan untuk hidup layak [1].

Klasterisasi atau Klasifikasi atau *Cluster* adalah salah satu cabang dari *Data Mining*. Menurut Athauda et al. (2009:201), beberapa pengertian tentang *Data Mining* adalah "*Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques*" (Gartner Group, 1995). Kemudian, "*Data mining is the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner*" (Hand et al., 2001). Juga, "*Data mining is an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases, and visualization to address the issue of information extraction from large data bases*" (Cabena et al., 1998). Dan juga, "*The extraction of interesting (non-trivial, implicit, previously unknown and potentially useful) patterns or knowledge from huge amount of data*" [2].

Menurut Cantu-Paz et al. (2002:50), *Data Mining* bersifat multi disiplin atau merupakan gabungan dari berbagai disiplin ilmu pengetahuan dan terapan, yang di definisikan sbb. : "*Data mining is a multidisciplinary field, borrowing ideas from machine learning and artificial intelligence, statistics, high performance computing, signal and image processing, mathematical optimization, pattern recognition, etc.*" Dan, *Data mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. *Data mining* adalah proses yang menggunakan metoda statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yg terkait dari berbagai database besar [3].

Pada algoritma *clustering*, data akan dikelompokkan menjadi *cluster-cluster* berdasarkan kemiripan satu data dengan yang lain. Prinsip dari *clustering* adalah memaksimalkan kesamaan antar anggota satu cluster dan meminimumkan kesamaan antar anggota cluster yang berbeda) [4]. *Cluster* atau cluster dapat diartikan kelompok, dengan demikian pada dasarnya analisis cluster akan menghasilkan sejumlah kluster (kelompok). Analisis ini diawali dengan pemahaman bahwa sejumlah data tertentu sebenarnya mempunyai kemiripan diantara anggotanya, karena itu dimungkinkan untuk mengelompokkan anggota-anggota yang mirip atau mempunyai karakteristik yang serupa tersebut dalam satu atau lebih dari satu kluster [5].

Analisis cluster akan membagi sejumlah data pada satu atau beberapa cluster tertentu. Pertanyaan yang timbul adalah "apa yang menjadi batas bahwa sejumlah data dapat disebut sebagai satu cluster ?". Secara logika, sebuah cluster yang baik adalah cluster yang mempunyai :

- a. *Homogenitas* (kesamaan) yang tinggi antara anggota dalam satu cluster (*with-in cluster*).
- b. *Heterogenitas* (perbedaan) yang tinggi antara cluster yang satu dengan cluster lainnya (*between cluster*).

Sehingga, *cluster* yang baik adalah *cluster* yang mempunyai anggota-anggota yang semirip mungkin satu dengan yang lain, namun sangat tidak mirip dengan anggota-anggota *cluster* yang lain. Disini, istilah "mirip" diartikan sebagai tingkat kesamaan karakteristik antara dua data [5].

Studi tentang pengklasteran terhadap data IPM juga di lakukan oleh beberapa penulis, yaitu M. Faturrahman yang menyoroti Klasifikasi Indeks Pembangunan Manusia (IPM) di Pulau Sumatera [6]. Juga Debby Armetiyana dkk yang membahas Pengklasteran IPM Metode Fuzzy C-Means [7]. Juga, Horainim Sibarani, dkk yang menggunakan Metode K-Means untuk Pengelompokkan IPM di Kabupaten/Kota Di Provinsi Sumatera Utara [8]. Dan, Fatkhurokhman Fauzi, dkk menulis tentang Klasifikasi IPM Kabupaten/Kota selindonesia dengan pendekatan smooth support vector machine (ssvm) kernel radial basis function (rbf) [9]. Algoritma K-Mean juga dipakai oleh Samuel & Akhmad Budi untuk meng-klaster IPM per-Kabupaten di Indonesia[10] .

2. Metode Penelitian

Analisis *clustering* pada dasarnya adalah mencari dan mengelompokkan data yang mirip satu dengan yang lain, maka kriteria mirip (*similarity*) adalah dasar dari metode *clustering*. Proses pengolahan data sehingga sekumpulan data mentah dapat dikelompokkan menjadi satu atau beberapa cluster adalah sbb. :

- 1) Menetapkan ukuran jarak antar-data. Mengukur kesamaan antar-obyek (*similarity*), secara prinsip dasar cluster yang mengelompokkan obyek yang mempunyai kemiripan, maka proses pertama adalah mengukur seberapa jauh ada kesamaan antar-obyek. Ada 3 metode yang digunakan :
 - a) Mengukur korelasi antara sepasang obyek pada beberapa variabel.
 - b) Mengukur jarak (*distance*) antara dua obyek.
 - c) Mengukur asosiasi antar-obyek.
- 2) Melakukan proses standarisasi data jika diperlukan.
- 3) Melakukan proses clustering. Setelah data yang dianggap mempunyai satuan yang satuan yang sangat berbeda diseragamkan & metode cluster ditentukan, misal. dipilih Euclidian, langkah selanjutnya adalah membuat cluster. Proses inti dari clustering adalah pengelompokan data, yang bisa dilakukan dengan 2 (dua) metode :
 - a) *Hierarchical Method*. Metode ini memulai pengelompokan dengan dua atau lebih obyek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke obyek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar-obyek. Dari yang paling mirip sampai paling tidak mirip. Semua obyek pada akhirnya hanya akan membentuk sebuah cluster. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut.
 - b) *Non-Hierarchical Method*. Berbeda dengan metode hierarki, metode ini justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (bisa 2, 3 atau lebih). Setelah jumlah cluster ditentukan, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan K-Means Cluster.
- 4) Melakukan penamaan cluster-cluster yang terbentuk
Setelah sejumlah cluster terbentuk, entah dengan metode hierarki atau non-hierarki, langkah selanjutnya adalah melakukan interpretasi terhadap cluster yang telah terbentuk, yaitu memberi nama spesifik untuk menggambarkan isi cluster tersebut.
- 5) Melakukan validasi dan profiling kluster

Kluster yang terbentuk kemudian diuji apakah hasil tersebut valid. Kemudian dilakukan proses profiling untuk menjelaskan karakteristik setiap cluster berdasar profil tertentu. [5].

3. Hasil dan Pembahasan

3.1. Analysis Missing Value & Replace Missing Value

Data yang diolah seluruhnya ada 515 yang terdiri atas 15 field atau variabel, yaitu nomor urut, nama Kota atau Kabupaten & skor Angka Indeks IPM tiap Kota atau Kabupaten, mulai tahun 2010 s/d 2022. Data disiapkan dalam format MS Excel 2019 (Tabel 1), yang nantinya akan dianalisis secara Deskriptif & Cluster. Data dalam format Excel, selanjutnya disesuaikan dengan format file Aplikasi SPSS (Tabel 2 & Tabel 3), diantaranya baris pertama akan otomatis menjadi nama field atau variabel pada Aplikasi SPSS. Seperti pada tampilan di bawah ini.

Tabel 1. Tabel Data format MS Excel

No	Kabupaten_Kota	Thn_2010	Thn_2011	Thn_2012	Thn_2013	Thn_2014	Thn_2015	Thn_2016	Thn_2017	Thn_2018	Thn_2019	Thn_2020	Thn_2021	Thn_2022
1	Banuwangi	60.60	61.83	61.25	61.88	62.18	63.16	63.82	64.41	64.74	65.70	66.03	66.41	67.27
2	Arabi Bagulu	62.36	63.13	64.23	64.87	65.27	66.05	66.96	67.37	68.02	68.91	68.94	69.22	69.82
3	Arabi Selatan	61.22	61.52	61.69	62.27	62.28	63.29	64.13	65.03	65.02	66.00	67.12	67.44	67.87
4	Arabi Tenggara	63.80	64.27	64.89	65.55	65.89	66.77	67.48	68.09	68.67	69.36	69.37	69.44	70.21
5	Arabi Timur	61.75	62.35	62.93	63.27	63.97	64.93	65.82	66.52	66.82	67.39	67.63	67.82	68.72
101	Yalimo	58.87	59.85	61.84	63.33	64.23	64.82	64.95	66.19	67.13	68.00	68.26	69.01	69.90
102	Puncak	61.44	61.68	61.85	62.73	62.83	63.41	63.98	64.88	65.01	65.70	65.03	65.17	65.87
103	Dugha	47.40	48.40	50.55	51.46	52.25	52.78	53.32	54.04	54.44	55.41	54.84	55.00	55.77
104	Intan Jaya	#VALUE!	48.07	41.80	42.69	43.51	44.35	44.82	45.68	46.55	47.51	47.79	48.34	49.25
105	Doyai	#VALUE!	46.12	46.94	47.74	48.12	48.38	48.50	49.07	49.35	50.11	49.46	49.06	50.42
106	Kota Jayapura	76.69	76.97	77.25	77.46	77.86	78.85	79.56	79.23	79.58	80.16	79.94	80.11	80.61

Tabel 2. Tabel Data format SPSS Data View

No	Kabupaten_Kota	Thn_2010	Thn_2011	Thn_2012	Thn_2013	Thn_2014	Thn_2015	Thn_2016	Thn_2017	Thn_2018	Thn_2019	Thn_2020	Thn_2021	Thn_2022
1	Banuwangi	60.60	61.83	61.25	61.88	62.18	63.16	63.82	64.41	64.74	65.70	66.03	66.41	67.27
2	Arabi Bagulu	62.36	63.13	64.23	64.87	65.27	66.05	66.96	67.37	68.02	68.91	68.94	69.22	69.82
3	Arabi Selatan	61.22	61.52	61.69	62.27	62.28	63.29	64.13	65.03	65.02	66.00	67.12	67.44	67.87
4	Arabi Tenggara	63.80	64.27	64.89	65.55	65.90	66.77	67.48	68.09	68.67	69.36	69.37	69.44	70.21
5	Arabi Timur	61.75	62.35	62.93	63.27	63.97	64.93	65.82	66.52	66.82	67.39	67.63	67.82	68.72
101	Yalimo	58.87	59.85	61.84	63.33	64.23	64.82	64.95	66.19	67.13	68.00	68.26	69.01	69.90
102	Puncak	61.44	61.68	61.85	62.73	62.83	63.41	63.98	64.88	65.01	65.70	65.03	65.17	65.87
103	Dugha	47.40	48.40	50.55	51.46	52.25	52.78	53.32	54.04	54.44	55.41	54.84	55.00	55.77
104	Intan Jaya	#VALUE!	48.07	41.80	42.69	43.51	44.35	44.82	45.68	46.55	47.51	47.79	48.34	49.25
105	Doyai	#VALUE!	46.12	46.94	47.74	48.12	48.38	48.50	49.07	49.35	50.11	49.46	49.06	50.42
106	Kota Jayapura	76.69	76.97	77.25	77.46	77.86	78.85	79.56	79.23	79.58	80.16	79.94	80.11	80.61

Tabel 3. Tabel Data format SPSS Variable View

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	No	String	154	0	None	None	3	Left	Nominal	Input
2	Kabupaten_Kota	String	26	0	None	None	14	Left	Nominal	Input
3	Thn_2010	Numeric	9	2	None	None	9	Right	Scale	Input
4	Thn_2011	Numeric	9	2	None	None	9	Right	Scale	Input
5	Thn_2012	Numeric	9	2	None	None	9	Right	Scale	Input
6	Thn_2013	Numeric	9	2	None	None	9	Right	Scale	Input
7	Thn_2014	Numeric	9	2	None	None	9	Right	Scale	Input
8	Thn_2015	Numeric	9	2	None	None	9	Right	Scale	Input
9	Thn_2016	Numeric	9	2	None	None	9	Right	Scale	Input
10	Thn_2017	Numeric	9	2	None	None	9	Right	Scale	Input
11	Thn_2018	Numeric	9	2	None	None	9	Right	Scale	Input
12	Thn_2019	Numeric	9	2	None	None	9	Right	Scale	Input
13	Thn_2020	Numeric	9	2	None	None	9	Right	Scale	Input
14	Thn_2021	Numeric	9	2	None	None	9	Right	Scale	Input
15	Thn_2022	Numeric	9	2	None	None	9	Right	Scale	Input

Sebelum proses cluster, di lakukan uji Missing Value Analysis atau pemeriksaan atas data yang kosong atau tidak ada. Ketidakadaan data ini disebabkan oleh keberadaan wilayah tsb baru di bentuk atau di dirikan. Dari Tabel 4, bisa di lihat bahwa ada Missing Data sebanyak 34 data (tahun 2010), 21 data (tahun 2011), 21 data (tahun 2012), dan masing-masing 4 data (dari tahun 2014 s/d tahun 2022).

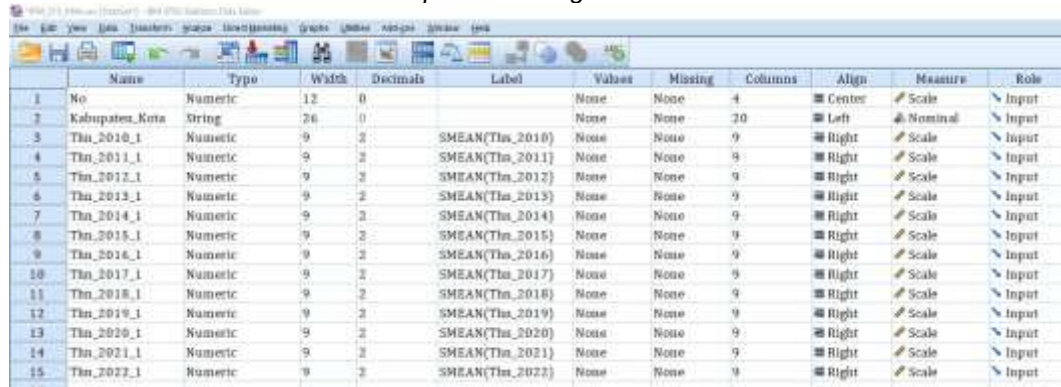
Tabel 4. Hasil Missing Value Analysis

	N	Mean	Std. Deviation	Missing		No. of Extremes ^a	
				Count	Percent	Low	High
Thn_2010	485	64.274866	7.0059427	34	6.6	12	5
Thn_2011	498	64.750924	7.1367079	21	4.0	16	5
Thn_2012	498	65.420040	7.0195519	21	4.0	16	6
Thn_2013	512	65.959121	6.9419635	7	1.3	16	6
Thn_2014	515	66.451942	6.8884171	4	.8	14	5
Thn_2015	515	67.095379	6.8713340	4	.8	16	7
Thn_2016	515	67.679282	6.8146213	4	.8	15	6
Thn_2017	515	68.248097	6.7005098	4	.8	17	10
Thn_2018	515	68.880524	6.6175840	4	.8	18	10
Thn_2019	515	69.530913	6.5437764	4	.8	16	10
Thn_2020	515	69.636971	6.5106848	4	.8	17	10
Thn_2021	515	69.932621	6.4920679	4	.8	17	10
Thn_2022	515	70.595204	6.4202271	4	.8	17	8

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

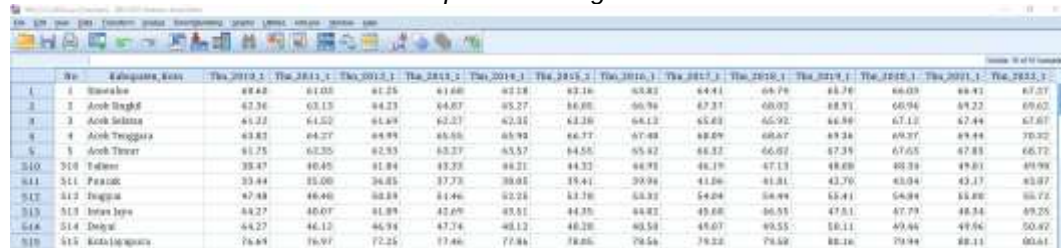
Proses berikutnya adalah mengisi data yang kosong tadi dengan data pengisi sesuai hasil Replace Missing Values seperti di Tabel 5 & Tabel 6 berikut ini. Dan, hasil Replace Missing Values inilah yang akan menjadi bahan untuk proses Cluster nantinya.

Tabel 5. Hasil Replace Missing Values - Variable View



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	No	Numeric	12	0		None	None	4	Center	Scale	Input
2	Kabupaten_Kota	String	26	0		None	None	20	Left	Nominal	Input
3	Thn_2010_1	Numeric	9	2	SMEAN(Thn_2010)	None	None	9	Right	Scale	Input
4	Thn_2011_1	Numeric	9	2	SMEAN(Thn_2011)	None	None	9	Right	Scale	Input
5	Thn_2012_1	Numeric	9	2	SMEAN(Thn_2012)	None	None	9	Right	Scale	Input
6	Thn_2013_1	Numeric	9	2	SMEAN(Thn_2013)	None	None	9	Right	Scale	Input
7	Thn_2014_1	Numeric	9	2	SMEAN(Thn_2014)	None	None	9	Right	Scale	Input
8	Thn_2015_1	Numeric	9	2	SMEAN(Thn_2015)	None	None	9	Right	Scale	Input
9	Thn_2016_1	Numeric	9	2	SMEAN(Thn_2016)	None	None	9	Right	Scale	Input
10	Thn_2017_1	Numeric	9	2	SMEAN(Thn_2017)	None	None	9	Right	Scale	Input
11	Thn_2018_1	Numeric	9	2	SMEAN(Thn_2018)	None	None	9	Right	Scale	Input
12	Thn_2019_1	Numeric	9	2	SMEAN(Thn_2019)	None	None	9	Right	Scale	Input
13	Thn_2020_1	Numeric	9	2	SMEAN(Thn_2020)	None	None	9	Right	Scale	Input
14	Thn_2021_1	Numeric	9	2	SMEAN(Thn_2021)	None	None	9	Right	Scale	Input
15	Thn_2022_1	Numeric	9	2	SMEAN(Thn_2022)	None	None	9	Right	Scale	Input

Tabel 6. Hasil Replace Missing Values - Data View



No	Kabupaten_Kota	Thn_2010_1	Thn_2011_1	Thn_2012_1	Thn_2013_1	Thn_2014_1	Thn_2015_1	Thn_2016_1	Thn_2017_1	Thn_2018_1	Thn_2019_1	Thn_2020_1	Thn_2021_1	Thn_2022_1
1	Banjar	61.62	61.05	61.25	61.68	62.18	62.18	62.82	64.41	65.78	66.03	66.91	67.57	68.27
2	Asah Tengah	62.36	63.13	64.23	64.87	65.27	66.02	66.78	67.31	68.02	68.71	69.94	69.22	69.62
3	Asah Selatan	61.23	61.52	61.89	62.23	62.35	63.38	64.12	65.43	65.92	66.88	67.12	67.44	67.87
4	Asah Tenggara	61.83	64.27	64.99	65.55	66.77	67.48	68.09	68.87	69.34	69.37	69.44	70.32	70.82
5	Asah Timur	61.25	62.35	62.93	63.37	63.57	64.55	65.42	66.32	66.82	67.39	67.65	67.83	68.72
10	Talensi	58.47	60.45	61.84	63.35	64.21	64.33	64.93	66.18	67.13	68.68	69.38	69.81	69.98
11	Pasau	53.44	55.00	56.85	57.73	58.05	59.41	59.94	61.06	61.81	63.70	63.94	63.17	63.87
12	Dugan	67.48	66.40	66.84	67.46	67.28	68.33	68.84	69.44	69.44	69.44	69.44	69.44	69.44
13	Intan Jaya	64.27	65.07	65.89	66.69	67.51	68.33	69.08	69.81	70.51	71.79	72.44	73.25	74.06
14	Dugan	64.27	65.07	65.89	66.69	67.51	68.33	69.08	69.81	70.51	71.79	72.44	73.25	74.06
15	Kota Jayapura	76.84	76.97	77.25	77.46	77.84	78.05	78.54	79.22	79.68	80.16	80.94	81.11	80.61

Dan, untuk memastikan bahwa data sudah lengkap, maka kita check lagi untuk Missing Values-nya. Berikut hasil Replace Missing Values yang mana data yang Missing sudah 0 semua. Berarti proses Cluster siap di laksanakan.

Tabel 7. Hasil Replace Missing Values yang sudah bebas Missing Value.

	N	Mean	Std. Deviation	Missing		No. of Extremes ^a	
				Count	Percent	Low	High
Thn_2010_1	515	64.2749	6.79842	0	.0	13	14
Thn_2011_1	515	64.7509	7.01770	0	.0	16	6
Thn_2012_1	515	65.4200	6.90249	0	.0	16	7
Thn_2013_1	515	65.9591	6.92168	0	.0	16	6
Thn_2014_1	515	66.4519	6.88842	0	.0	14	5
Thn_2015_1	515	67.0954	6.87133	0	.0	16	7
Thn_2016_1	515	67.6793	6.81462	0	.0	15	6
Thn_2017_1	515	68.2481	6.70051	0	.0	17	10
Thn_2018_1	515	68.8805	6.61758	0	.0	18	10
Thn_2019_1	515	69.5309	6.54378	0	.0	16	10
Thn_2020_1	515	69.6370	6.51068	0	.0	17	10
Thn_2021_1	515	69.9326	6.49207	0	.0	17	10
Thn_2022_1	515	70.5952	6.42023	0	.0	17	8

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

Tabel 8. Rekapitan Hasil Replace Missing Values yang sudah bebas Missing Value.

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
515	100.0	0	.0	515	100.0

a. Average Linkage (Between Groups)

3.2. Hasil Proses Cluster & Profiling Cluster

Hasil dari proses Cluster adalah menampilkan angka-angka Cluster pada setiap wilayah, seperti pada Tabel 9 berikut ini. Angka-angka pada setiap kolom menunjukkan posisi cluster masing-masing sesuai dengan pembagian jumlah cluster yang telah di tentukan. Tabel 9 ini hanya menampilkan 10 baris pertama & 10 baris terakhir dari 515 baris data seluruhnya.

Tabel 9. Hasil Proses Cluster dari SPSS (Cluster Membership)

Cluster Membership						
Case	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1:Simelue	1	1	1	1	1	1
2:Aceh Singkil	1	1	1	1	1	1
3:Aceh Selatan	1	1	1	1	1	1
4:Aceh Tenggara	1	1	1	1	1	1
5:Aceh Timur	1	1	1	1	1	1
6:Aceh Tengah	2	1	1	1	1	1
7:Aceh Barat	1	1	1	1	1	1
8:Aceh Besar	2	1	1	1	1	1
9:Pidie	1	1	1	1	1	1
10:Bireuen	2	1	1	1	1	1
506:Mamberamo Raya	6	5	4	3	2	2
507:Nduga	7	6	5	4	3	2
508:Lanny Jaya	6	5	4	3	2	2
509:Mamberamo Tengah	6	5	4	3	2	2
510:Yalimo	6	5	4	3	2	2
511:Puncak	6	5	4	3	2	2
512:Dogiyai	4	3	3	1	1	1
513:Intan Jaya	5	4	4	3	2	2
514:Diyai	5	4	4	3	2	2
515:Kota Jayapura	3	2	2	2	1	1

Selanjutnya, adalah proses profiling cluster, yaitu dengan sortasi berdasarkan cluster masing-masing, dengan kunci sortasi seperti di Tabel 10 berikut ini.

Tabel 10. Proses Profiling Cluster dengan Sortasi secara Ascending

Sort

☒ My data has headers

Column	Sort On	Order
Sort by Rata-Rata Per Baris	Cell Values	Smallest to Largest
Then by Thn_2010_1	Cell Values	Smallest to Largest
Then by Thn_2011_1	Cell Values	Smallest to Largest
Then by Thn_2012_1	Cell Values	Smallest to Largest
Then by Thn_2013_1	Cell Values	Smallest to Largest
Then by Thn_2014_1	Cell Values	Smallest to Largest
Then by Thn_2015_1	Cell Values	Smallest to Largest
Then by Thn_2016_1	Cell Values	Smallest to Largest
Then by Thn_2017_1	Cell Values	Smallest to Largest
Then by Thn_2018_1	Cell Values	Smallest to Largest
Then by Thn_2019_1	Cell Values	Smallest to Largest
Then by Thn_2020_1	Cell Values	Smallest to Largest
Then by Thn_2021_1	Cell Values	Smallest to Largest
Then by Thn_2022_1	Cell Values	Smallest to Largest

Tabel 11. Tampilan 10 data teratas dari hasil Cluster

Cluster Membership						
Case	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1:Simelue	1	1	1	1	1	1
2:Aceh Singkil	1	1	1	1	1	1
3:Aceh Selatan	1	1	1	1	1	1
4:Aceh Tenggara	1	1	1	1	1	1
5:Aceh Timur	1	1	1	1	1	1
7:Aceh Barat	1	1	1	1	1	1
9:Pidie	1	1	1	1	1	1
11:Aceh Utara	1	1	1	1	1	1
12:Aceh Barat Daya	1	1	1	1	1	1
13:Gaya Lues	1	1	1	1	1	1

Tabel 12. Tampilan 10 data terendah dari hasil Cluster

Cluster Membership						
Case	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters	2 Clusters
498:Asmat	6	5	4	3	2	2
499:Yahukimo	6	5	4	3	2	2
500:Pegunungan Bintang	6	5	4	3	2	2
501:Tolikara	6	5	4	3	2	2
506:Mamberamo Raya	6	5	4	3	2	2
508:Lanny Jaya	6	5	4	3	2	2
509:Mamberamo Tengah	6	5	4	3	2	2
510:Yalimo	6	5	4	3	2	2
511:Puncak	6	5	4	3	2	2
507:Nduga	7	6	5	4	3	2

Tabel 13. Tampilan 10 data terendah dari hasil Rata-rata angka IPM tiap Tahun

No	Kabupaten_Kota	Thn_2010_1	Thn_2011_1	Thn_2012_1	Thn_2013_1	Thn_2014_1	Thn_2015_1	Thn_2016_1	Thn_2017_1	Thn_2018_1	Thn_2019_1	Thn_2020_1	Thn_2021_1	Thn_2022_1	Rata-Rata Per Baris
501	Tolikara	43.44	44.41	44.86	45.68	46.16	46.38	47.11	47.89	48.85	49.68	49.50	49.60	50.51	47.24
499	Yahukimo	40.63	41.72	43.82	45.63	46.36	46.63	47.13	47.95	48.51	49.25	49.37	49.48	50.25	46.67
513	Intan Jawa	64.27	40.07	41.89	42.69	43.51	44.35	44.82	45.68	46.55	47.51	47.79	48.34	49.25	46.67
508	Lanny Jaya	41.49	41.90	42.53	43.05	43.28	44.18	45.16	46.49	47.34	48.00	47.86	48.68	49.62	45.35
494	Puncak Jaya	38.83	40.36	41.85	43.36	44.32	44.87	45.49	46.57	47.39	48.33	48.37	48.99	49.84	45.27
510	Yalimo	38.47	40.45	41.84	43.33	44.21	44.32	44.95	46.19	47.13	48.08	48.34	49.01	49.90	45.09
509	Mamberamo Tengah	39.37	40.17	41.39	42.43	43.19	43.55	44.15	45.50	46.41	47.23	47.57	48.32	49.25	44.50
500	Pegunungan Bintang	35.45	36.61	37.82	38.94	39.68	40.91	41.90	43.24	44.22	45.21	45.44	46.28	47.21	41.76
511	Puncak	33.44	35.08	36.85	37.73	38.05	39.41	39.96	41.06	41.81	42.70	43.04	43.17	43.87	39.71
507	Nduga	19.62	21.12	23.07	24.42	25.38	25.47	26.56	27.87	29.42	30.75	31.55	32.84	34.10	27.09

Tabel 14. Tampilan 10 data tertinggi dari hasil Rata-rata angka IPM tiap Tahun

No	Kabupaten_Kota	Thn_2010_1	Thn_2011_1	Thn_2012_1	Thn_2013_1	Thn_2014_1	Thn_2015_1	Thn_2016_1	Thn_2017_1	Thn_2018_1	Thn_2019_1	Thn_2020_1	Thn_2021_1	Thn_2022_1	Rata-Rata Per Baris
228	Kota Yogyakarta	82.72	82.98	83.29	83.61	83.78	84.56	85.32	85.49	86.11	86.65	86.61	87.18	87.69	85.08
19	Kota Banda Aceh	80.26	80.87	81.30	81.84	82.22	83.25	83.73	83.95	84.37	85.07	85.41	85.71	86.28	83.41
157	Kota Jakarta Selatan	80.26	81.22	81.72	82.72	82.94	83.37	83.94	84.13	84.44	84.75	84.72	84.90	85.21	83.41
283	Kota Denpasar	79.19	79.77	80.45	81.32	81.65	82.24	82.58	83.01	83.30	83.68	83.93	84.03	84.37	82.27
227	Sleman	79.69	80.04	80.10	80.26	80.73	81.20	82.15	82.85	83.42	83.85	83.84	84.00	84.31	82.03
439	Kota Kendari	78.13	79.43	79.97	80.91	81.30	81.43	81.66	81.83	82.22	82.86	83.53	84.15	84.51	81.69
220	Kota Salatiga	78.35	78.76	79.10	79.37	79.98	80.96	81.14	81.68	82.41	83.12	83.14	83.60	84.35	81.23
158	Kota Jakarta Timur	78.06	78.82	79.52	79.88	80.40	80.73	81.28	81.61	82.06	82.69	82.66	82.97	83.45	81.09
69	Kota Padang	78.44	78.68	79.00	79.23	79.83	80.36	81.06	81.58	82.25	82.68	82.82	82.90	83.29	80.93
221	Kota Semarang	76.96	77.58	78.04	78.68	79.24	80.23	81.19	82.01	82.72	83.19	83.05	83.55	84.08	80.81

Berikutnya adalah menghitung jumlah anggota masing-masing cluster & membuat pelevelan Profiling Data setiap Cluster terkait Angka IPM. Untuk hasil 7 cluster, bisa di rinci anggota masing-masing Cluster adalah sbb. :

- Anggota dari Cluster 1 sebanyak 317 Kabupaten/Kota dengan level angka IPM-nya “Sangat Sangat Tinggi”.
- Anggota dari Cluster 2 sebanyak 115 Kabupaten/Kota dengan level angka IPM-nya “Sangat Tinggi”.
- Anggota dari Cluster 3 sebanyak 60 Kabupaten/Kota dengan level angka IPM-nya “Tinggi”.
- Anggota dari Cluster 4 sebanyak 9 Kabupaten/Kota dengan level angka IPM-nya “Sedang”.
- Anggota dari Cluster 5 sebanyak 3 Kabupaten/Kota dengan level angka IPM-nya “Rendah”.
- Anggota dari Cluster 6 sebanyak 10 Kabupaten/Kota dengan level angka IPM-nya “Sangat Rendah”.
- Anggota dari Cluster 7 sebanyak 1 Kabupaten/Kota dengan level angka IPM-nya “Sangat Sangat Rendah”.

Tabel 15. Tampilan 7 cluster dengan rincian data tiap clusternya dari cluster 1 s/d 7

Jumlah Cluster	Data Tiap Cluster	Level Profiling Angka IPM
7 Clusters	Cluster 1 317	Sangat Sangat Tinggi
	Cluster 2 115	Sangat Tinggi
	Cluster 3 60	Tinggi
	Cluster 4 9	Sedang
	Cluster 5 3	Rendah
	Cluster 6 10	Sangat Rendah
	Cluster 7 1	Sangat Sangat Rendah
Jumlah Data	515	

Demikian juga untuk data tiap Cluster berikutnya yaitu untuk 6 & 5 cluster tertampil di Tabel di bawah ini.

Tabel 16. Tampilan 6 cluster dengan rincian data tiap clusternya dari cluster 1 s/d 6

Jumlah Cluster	Data Tiap Cluster	Level Profiling Angka IPM
6 Clusters	Cluster 1 432	Sangat Sangat Tinggi
	Cluster 2 60	Sangat Tinggi
	Cluster 3 9	Tinggi
	Cluster 4 3	Rendah
	Cluster 5 10	Sangat Rendah
	Cluster 6 1	Sangat Sangat Rendah
Jumlah Data	515	

Tabel 17. Tampilan 5 cluster dengan rincian data tiap clusternya dari cluster 1 s/d 5

Jumlah Cluster	Data Tiap Cluster	Level Profiling Angka IPM
5 Clusters	Cluster 1 432	Sangat Tinggi
	Cluster 2 60	Tinggi
	Cluster 3 9	Sedang
	Cluster 4 13	Rendah
	Cluster 5 1	Sangat Rendah
Jumlah Data	515	

4.Kesimpulan

Angka IPM sebagai indikator pembangunan bisa di maknai secara lebih jelas dengan menggunakan analisa cluster. Salah satu metode yang bisa di gunakan adalah Cluster Hierarki. Metode Cluster Hierarki di pilih dengan maksud bisa memberikan gambaran atas pilihan jumlah cluster. Pembaca di harapkan bisa menentukan pilihan jumlah cluster yang berapa cluster yang akan di pakai untuk mempelajari angka IPM ini.

Ucapan Terimakasih

Terima kasih kepada Fakultas Teknik Universitas Suryakancana yang telah menjadi wadah bagi para peneliti untuk mengembangkan penelitian jurnal ini. Semoga penelitian ini dapat memberikan kontribusi yang besar bagi kemajuan ilmu pengetahuan di Indonesia.

Referensi

- [1] [1<https://jambi.bps.go.id/subjec...> n.d.]
- [2] Athauda, Rukshan., Tissera, Menik., and Fernando, Chandrika., Data Mining Applications: Promise and Challenges. Publikasi pada Data Mining and Knowledge Discovery in Real Life Applications, Julio Ponce and Adem Karahoca, 2009
- [3] Han, Jiawei. & Kamber, Micheline., Data Mining: Concepts and Techniques Second Edition, University of Illinois at Urbana-Champaign, 2006.
- [4] Sumiah, M. "Model Data Mining Pada Sistem Informasi Akademik: Studi Kasus Universitas Kuningan." Model Data Mining Pada Sistem Informasi Akademik: Studi Kasus Universitas Kuningan (2011).
- [5] Santoso, Singgih., Statistik Multivariat, Konsep dan Aplikasi dengan SPSS, Elex Media Komputindo, 2010
- [6] Fathurrahman, M., & Qisthi, N. (2021). Klasifikasi Indeks Pembangunan Manusia (Ipm) Di Pulau Sumatera Pada Dataset Multi-Class Dengan Metode Artificial Neural Network (Ann). In Prosiding Seminar Nasional Fisika (Vol. 7, Issue 0).
- [7] Margaretta, Debby Armetiyana, Pengklasteran Provinsi-Provinsi Di Indonesia Berdasarkan Indikator Yang Mempengaruhi Indeks Pembangunan Manusia Menggunakan Metode Fuzzy C-Means, Izzati Rahmi Hg * , Hazmira Yozza Studi P Jurnal Matematika Unand (2021) 10(1) 79-86
- [8] Sibarani, H., Saputra, W., Gunawan, I., Masruro Nasution Program Studi Teknik Informatika Stikom Tunas Bangsa Pematangsiantar, Z., & Jendral Sudirman Blok, J. A. (2022). Penerapan Metode K-Means Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia. In Jurnal Mahasiswa Teknik Informatika (Vol. 6, Issue 1).
- [9] Fatkhurokhman Fauzi, Moh. Yamin Darsyah Dan Tiani Wahyu Utami, Seminar Nasional Pendidikan, Sains Dan Teknologi Klasifikasi Indeks Pembangunan Manusia Kabupaten/Kota Se-Indonesia Dengan Pendekatan Smooth Support Vector Machine (Ssvm) Kernel Radial Basis Function (Rbf)
- [10] Samuel Dan Akhmad Budi, Klasterisasi Indeks Pembangunan Manusia (Ipm) Per Kabupaten Di Indonesia Dengan Menggunakan Algoritma K-Means, Jurnal Informatika Dan Bisnis.