

Klasterisasi Kabupaten/Kota di Provinsi Kalimantan Barat Berdasarkan Rasio Guru dengan Murid di Tingkat SD, SMA, dan SMA Serta IPM Tahun 2022 Menggunakan Metode K-Means Clustering

District/City Clustering in West Kalimantan Province Based on Teacher to Student Ratios at Elementary, Senior High School, and Senior High School Levels and HDI in 2022 Using the K-Means Clustering Method

Mufrih Nur Huda Tri Putra^{1*}

¹Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Gadjah Mada, Bulaksumur, Yogyakarta, 55281, Indonesia;

*Penulis korespondensi. *e-mail:* e-mail: Mufrih.n.h@mail.ugm.ac.id
(Diterima: 27 Juni 2023; Disetujui: 31 Juli 2023)

ABSTRACT

The provincial government of West Kalimantan wants to accelerate HDI growth so that through this study it can classify regencies and cities that need more attention in the development of education in their area. The purpose of writing this journal is to classify regencies and cities in West Kalimantan Province according to the ratio of teachers and students at all levels of elementary, junior high, and high school as well as the Human Development Index (IPM) in 2022 so that it can be considered in formulating policies to accelerate the increase in HDI from the dimension knowledge. The data used in this study is data downloaded from the West Kalimantan Province BPS website. Data were analyzed using the cluster analysis method with the K-means clustering method to form several clusters. The results obtained formed three clusters, namely clusters 0, 1, and 2. Cluster 0 has the characteristics of a high SD ratio, a high SMP ratio, a moderate SMA ratio, and a high HDI. Cluster 1 with the characteristics of a low primary school ratio, low junior high school ratio, low high school ratio, and low HDI. Cluster 2 has the characteristics of a moderate SD ratio, a moderate SMP ratio, a high SMA ratio, and a moderate HDI.

Keywords: cluster, education, HDI, k-means, teacher ratio

ABSTRAK

Pemerintah provinsi Kalimantan Barat ingin mempercepat pertumbuhan IPM sehingga melalui kajian ini dapat mengklasifikasikan kabupaten dan kota yang harus lebih diperhatikan dalam pembangunan pendidikan di daerahnya. Tujuan penulisan jurnal ini adalah untuk mengklasifikasikan kabupaten dan kota di Provinsi Kalimantan Barat menurut rasio guru dan murid pada semua jenjang SD, SMP, dan SMA serta Indeks Pembangunan Manusia (IPM) tahun 2022, sehingga dapat menjadi pertimbangan dalam perumusan kebijakan percepatan peningkatan IPM dari dimensi pengetahuan. Data yang digunakan dalam penelitian ini adalah data yang diunduh dari laman BPS Provinsi Kalimantan Barat. Data dianalisis menggunakan metode analisis kluster dengan metode pengklasteran K-means untuk membentuk beberapa kluster. Hasil yang diperoleh terbentuk tiga kluster, yaitu kluster 0, 1, dan 2. Kluster 0 dengan karakteristik rasio SD tinggi, rasio SMP tinggi, rasio SMU sedang, dan IPM tinggi.

Klaster 1 dengan karakteristik rasio SD rendah, rasio SMP rendah, rasio SMU rendah, dan IPM rendah. Klaster 2 dengan karakteristik rasio SD sedang, rasio SMP sedang, rasio SMU tinggi, dan IPM sedang.

Kata kunci: klaster, pendidikan, IPM, *k-means*, rasio guru

PENDAHULUAN

Indeks Pembangunan Manusia menjadi satu di antara tolak ukur dalam keberhasilan upaya membangun kualitas hidup manusia di suatu daerah. IPM Provinsi Kalimantan Barat sendiri berada di bawah nilai IPM nasional berada di peringkat 30 dari 34 Provinsi. Pembentukan IPM mengandung tiga dimensi dasar yaitu umur panjang dan hidup sehat, pengetahuan dan taraf hidup yang layak. Dimensi pengetahuan dapat dilihat melalui unsur-unsur pendidikan, seperti peserta didik, tenaga pendidik, dan lingkungan pendidikan yang harus memadai. Pendidikan yang baik memberikan dampak positif bagi kehidupan masyarakat. Percepatan peningkatan IPM dapat diraih dengan memperbaiki dimensi-dimensi yang berkaitan dengan IPM itu sendiri. Pemanfaatan data yang tersedia dengan pengolahan dan metode analisis yang tepat dapat menjadi pertimbangan yang baik dalam pengambilan kebijakan.

Pemerataan pendidikan perlu dilakukan agar seluruh lapisan masyarakat memiliki kesempatan yang sama dalam memperoleh pendidikan. Menurut Marita Qori'atunnadyah (2022:34) menyatakan bahwa pentingnya peranan guru dalam proses belajar mengajar. Saat proses pembelajaran terdapat batasan jumlah tertentu antara guru dan jumlah peserta didik. Perbandingan jumlah guru dan murid disebut rasio guru murid.

Semakin kecil rasio guru-murid, maka pengawasan guru terhadap murid semakin tinggi, yang berarti efek dan efisiensi proses pembelajaran dapat ditingkatkan, dan berdampak baik pada peningkatan kualitas pembelajaran. Misalnya, rasio guru murid di pendidikan dasar adalah 1:15, artinya setiap 1 guru dapat mengajar hingga 15 siswa sekolah dasar. Pasal 17 Peraturan Pemerintah Nomor 74 Tahun 2008 tentang Guru menyatakan bahwa pada tingkat SD, SMP, dan SMA sebaiknya satu orang guru membawahi 20 murid. Sedangkan di tingkat SMK, idealnya satu guru bertanggung jawab atas 15 siswa (SIMPUH KEMENAG, 2020).

Klasterisasi merupakan metode penganalisaan data, yang sering kali dimasukkan sebagai salah satu metode *Data Mining*, dengan tujuan utama yaitu untuk mengelompokkan objek-objek data berdasarkan karakteristik yang dimiliki. Pengelompokan objek didasari pada kemiripan ciri-ciri umum antar objek, yang mana objek yang ada pada setiap kelompok atau klaster akan memiliki kemiripan satu sama lain. Analisis klaster biasa digunakan pada berbagai bidang ilmu seperti psikologi, sosiologi, biologi, ekonomi, bisnis, dan lain-lain. Oleh karena itu, pembentukan klaster berdasarkan rasio guru dengan murid di tingkat SD, SMP, dan SMU serta IPM tahun 2022 dapat menentukan daerah yang perlu diperhatikan oleh pemerintah Provinsi Kalimantan Barat, terutama dalam hal rasio guru dan murid sehingga selaras dengan tujuan percepatan peningkatan IPM di provinsi Kalimantan Barat. Metode klasterisasi yang digunakan adalah analisis klaster *K-means*. Pemilihan metode *K-means* karena metode ini memiliki keunggulan dapat mengelompokkan objek besar dan *outlier* dengan sangat cepat sehingga dapat menangani pengelompokan tersebut, menjadikannya ideal untuk digunakan pada data yang dimiliki. Hal ini didasarkan pada fakta lapangan bahwa jumlah siswa yang bersekolah di kota dan kabupaten cenderung berbeda jauh jumlahnya sehingga kemungkinan terdapat data pencilan.

METODOLOGI

Data penelitian yang digunakan merupakan data informasi pendidikan pada laman BPS Provinsi Kalimantan Barat. Variabel-variabel pada penelitian ini adalah jumlah rasio guru dengan murid berdasarkan tingkatan pendidikan, yaitu SD, SMP, dan SMA serta IPM tahun 2022.

Tabel 1. Data Rasio Murid dengan Guru di Tingkat SD, SMP, dan SMA serta IPM tahun 2022.

Daerah	IPM	Rasio SD	Rasio SMP	Rasio SMU
(1)	(2)	(3)	(4)	(5)
Sambas	67,95	16,88	15,82	17,17
Bengkayang	68,74	13,96	15,08	14,30
Landak	67,17	13,87	14,03	18,16
Mempawah	66,94	13,86	12,46	16,71
Sanggau	66,91	13,75	14,98	19,08
Ketapang	67,92	12,87	11,06	16,59
Sintang	67,86	14,48	13,63	15,02
Kapuas Hulu	66,70	9,83	9,29	14,35
Sekadau	65,58	15,22	15,3	17,21
Melawi	66,81	10,21	9,26	15,05
Kayong Utara	63,81	11,37	10,77	15,08
Kubu Raya	68,91	15,02	14,14	14,2
Kota Pontianak	80,48	20,11	17,84	16,45
Kota Singkawang	72,89	17,87	16,28	14,81

Sumber : *kalbar.bps.go.id* (diolah)

Data yang diperoleh akan dianalisis dengan metode klasterisasi *K-means* menggunakan program python. Ada beberapa langkah *data mining* yang dijalankan pada penelitian ini, sebagai berikut:

1. *Selection*

Pemilihan variabel yang akan digunakan sehingga tidak terdapat informasi berulang yang tidak dibutuhkan dalam proses penelitian.

2. *Preprocessing*

Penggabungan variabel-variabel yang sebelumnya berasal dari *dataset* yang berbeda guna mempermudah analisis.

3. *Transformation*

Pengubahan format data sesuai dengan pengolahan *data mining* agar mengoptimalkan model.

4. *Data Mining*

Penggunaan metode untuk mendapatkan pengetahuan baru dari data yang telah diproses.

5. *Evaluation/Interpretation*

Pengidentifikasian pola-pola yang menarik sebagai hasil dari analisis yang telah dilakukan sebelumnya dan melihat performa dari model.

6. *Knowledge*

Pengetahuan yang diperoleh dapat disebarluaskan kepada pemangku jabatan sehingga dapat menjadi pertimbangan dalam pengambilan keputusan.

Analisis clustering *K-Means* merupakan salah satu algoritma yang dapat mengelompokkan data heterogen dalam teknologi data mining, karena algoritma *clustering* pada dasarnya hanya dapat mengidentifikasi nilai atribut yang homogen. Algoritma *K-Means Clustering* akan memilih sebanyak *k* titik awal *centroid* secara acak dan melakukan perhitungan jarak masing-masing observasi ke masing-masing *centroid* tersebut dengan menggunakan *Euclidean Distance*. Dari perhitungan jarak tersebut, akan diambil jarak yang paling dekat dari masing-masing observasi untuk menentukan masuk ke *cluster* manakah observasi tersebut.

Misalnya, diberikan sekumpulan objek $x = (x_1, x_2, \dots, x_n)$, algoritma pengelompokan *K-means* membagi x menjadi *k cluster*, masing-masing dengan *centroid*. Awalnya, algoritma analisis pengelompokan *K-Means* secara acak memilih *k* objek sebagai *centroid*, dan kemudian menggunakan jarak Euclidean untuk menghitung jarak antara objek dan *centroid*, menempatkan objek di *cluster* terdekat yang dihitung dari titik tengah *cluster*. Menentukan *centroid* jika semua objek telah ditempatkan di *cluster* terdekat. Penentuan *centroid* dan penempatan objek pada *cluster* dilakukan

berulang kali hingga nilai *centroid* semua *cluster* tidak berubah lagi. Secara umum, metode analisis clustering *K-Means* menggunakan algoritma berikut:

1. Uji asumsi *k-means clustering*, yaitu uji asumsi multikolinearitas dengan menggunakan nilai VIF dan kecukupan sampel. Asumsi multikolinearitas terpenuhi saat nilai $VIF < 10$ (Sarwoko dalam Zaenuddin, 2015). Rumus yang digunakan dalam perhitungan nilai VIF adalah, sebagai berikut:

$$VIF_i = \frac{1}{1-R_i^2} \quad (1)$$

2. Pemilihan jumlah klaster yang optimal dengan beberapa metode. Metode yang digunakan pada penelitian ini adalah metode *silhouette*.
3. Metode *silhouette* dapat melihat kualitas dan intensitas *cluster*, atau seberapa baik objek ditempatkan di dalam *cluster* tertentu. Metode ini adalah gabungan dari metode separasi dan kohesi. Rumus yang digunakan nilai *silhouette*, sebagai berikut:

$$s(x_i) = \frac{b(x_i) - a(x_i)}{\max\{b(x_i), a(x_i)\}} \quad (2)$$

Dengan keterangan, sebagai berikut:

x_i : elemen dalam klaster π_k ;

$a(x_i)$: rata-rata jarak dari x_i ke setiap elemen lain dalam klaster π_k ;

$b(x_i)$: rata-rata jarak terkecil dari x_i ke setiap titik dalam klaster π_k untuk $l \neq k$.

4. Pembentukan titik pusat klaster (*centroid*) secara random sebanyak jumlah klaster optimal. Rumus yang digunakan dalam menghitung *centroid*, sebagai berikut:

$$v = \frac{\sum_{i=1}^n x_i}{n}, i = 1, 2, 3, \dots, n \quad (3)$$

dengan,

v : *centroid* pada klaster;

x_i : observasi ke- i ;

n : jumlah anggota klaster.

5. Perhitungan jarak antara observasi dengan *centroid* menggunakan *euclidean distance* dengan formula, sebagai berikut:

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, i = 1, 2, 3, \dots, n \quad (4)$$

dengan,

x_i : observasi x ke- i ;

y_i : observasi y ke- i ;

n : banyaknya observasi.

6. Pengalokasian tiap data observasi ke dalam klaster dengan jarak *centroid* terdekat secara iteratif.
7. Pembentukan *centroid* baru dengan rumus yang digunakan untuk membentuk *centroid* sebelumnya.
8. Iterasi langkah ketiga dan seterusnya hingga diperoleh posisi *centroid* yang konsisten.

HASIL DAN PEMBAHASAN

Analisis Data Eksploratif

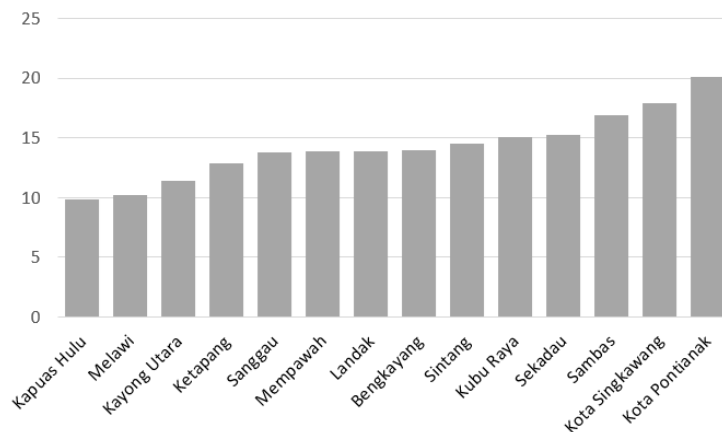
Tahapan pertama yang dilakukan sebelum melakukan analisis adalah eksplorasi data. Hal ini dilakukan dengan tujuan mengetahui gambaran data secara umum. Eksplorasi data dilakukan dengan membentuk diagram batang untuk variabel rasio SD, rasio SMP, rasio SMU, dan IPM.

Tabel 2. Statistika Deskriptif

Variabel	IPM	Rasio SD	Rasio SMP	Rasio SMU
(1)	(2)	(3)	(4)	(5)
Mean	68,47	14,24	13,57	16,01
Min	63,81	9,83	9,26	14,86
Max	80,48	20,11	17,84	19,08

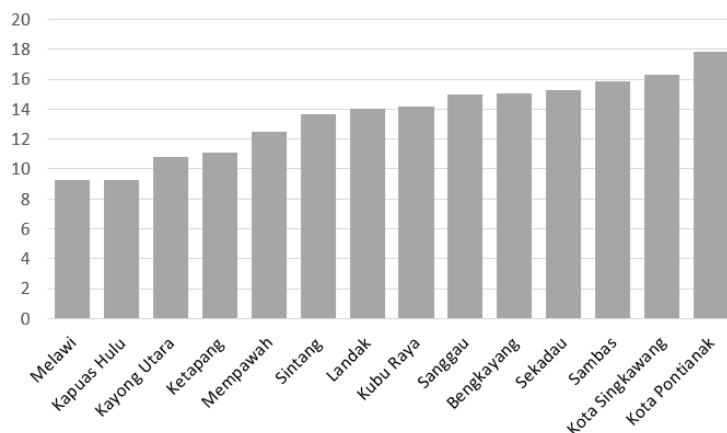
Sumber : *kalbar.bps.go.id* (diolah)

Rata-rata IPM seluruh daerah di Kalimantan Barat adalah 68.47. Rata-rata rasio SD, SMP, dan SMU secara berurutan adalah 14.24, 13.57, dan 16.01. Berdasarkan diagram batang yang terbentuk untuk variabel rasio SD diperoleh informasi bahwa Kapuas Hulu menjadi kabupaten dengan rasio SD terendah dan Kota Pontianak menjadi daerah dengan rasio SD tertinggi. Diagram batang untuk variabel rasio SMP menunjukkan bahwa Melawi menjadi kabupaten dengan rasio SMP terendah dan Kota Pontianak kembali menjadi daerah dengan rasio SMP tertinggi. Diagram batang untuk variabel rasio SMU menunjukkan bahwa Kubu Raya menjadi kabupaten dengan rasio SMU terendah dan Sanggau menjadi daerah dengan rasio SMU tertinggi. Selain itu, Diagram batang untuk variabel IPM menunjukkan Kubu Raya menjadi daerah dengan IPM terendah dan Sanggau menjadi daerah dengan IPM tertinggi.



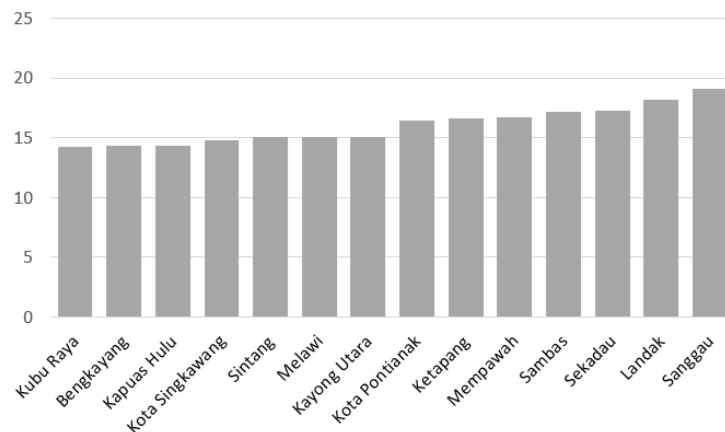
Gambar 1. Rasio Guru dengan Murid Tingkat SD

Sumber : *kalbar.bps.go.id* (diolah)



Gambar 2. Rasio Guru dengan Murid Tingkat SMP

Sumber : *kalbar.bps.go.id* (diolah)



Gambar 3. Rasio Guru dengan Murid Tingkat SMA
Sumber : *kalbar.bps.go.id* (diolah)

Uji Asumsi Multikolinearitas dan Kecukupan Sampel

Kecukupan sampel dimaksudkan agar sampel yang dimiliki dapat merepresentasikan populasi. Penelitian ini menggunakan data populasi itu sendiri, yaitu kabupaten dan kota di Provinsi Kalimantan Barat sehingga asumsi kecukupan sampel secara otomatis terpenuhi. Selanjutnya dilakukan uji multikolinearitas dengan melihat nilai *variance inflation factor* (VIF).

Tabel 3. Nilai VIF Sebelum Transformasi

Variabel	VIF
(1)	(2)
Rasio SD	287,94
Rasio SMP	263,79
Rasio SMU	100,59
IPM	147,58

Sumber : *kalbar.bps.go.id* (diolah)

Berdasarkan nilai VIF yang bernilai lebih dari 10, maka disimpulkan bahwa terdapat multikolinearitas antar variabel yang ada. Terdapat beberapa metode untuk mengatasi multikolinearitas yaitu transformasi log natural, *partial least squares* (PLS), metode *principal component analysis* (PCA) dan metode *stepwise* (Marcus et al., 2012; Pendi, 2021; D.N.P. Sari & Sukestiyarno, 2021) ; Silitonga et al., 2021; Supriyadi et al., 2017). Selanjutnya, akan dilakukan transformasi data dengan fungsi logaritma natural. Setelah dilakukan transformasi logaritma natural diperoleh nilai VIF, sebagai berikut:

Tabel 4. Nilai VIF Setelah Transformasi

Variabel	VIF
(1)	(2)
Rasio SD	7,76
Rasio SMP	7,61
Rasio SMU	6,92
IPM	7,14

Sumber : *kalbar.bps.go.id* (diolah)

Selanjutnya, data yang akan digunakan adalah data yang telah ditransformasi dengan fungsi logaritma natural.

Penentuan Jumlah Kluster Optimal

Penentuan jumlah *cluster* yang paling optimal dapat dilakukan dengan validasi hasil *cluster*. Pada penelitian ini, validasi dilakukan dengan menggunakan metode validasi *Silhouette Index*.

Tabel 5. *Silhouette Score* dari Beberapa Jumlah Cluster

Jumlah Kluster	Silhouetter Score
(1)	(2)
3	0.396
4	0.343
5	0.327
6	0.346
7	0.363
8	0.377

Sumber : *kalbar.bps.go.id* (diolah)

Dari tabel tersebut didapatkan informasi, nilai koefisien *Silhouette* terbesar yang berada pada antara jumlah kluster tiga sampai dengan delapan adalah kluster tiga. Kluster yang akan digunakan dalam metode klaterisasi *K-Means* ini adalah tiga. *Silhouetter score* yang diperoleh untuk jumlah kluster tiga adalah 0.396 sehingga pada analisis kluster akan dibentuk tiga kluster.

Centroid Data

Dalam penerapan algoritma K-Medoid nilai *centroid* dapat ditentukan secara acak dari data yang didapat dengan ketentuan bahwa pengelompokkan yang diinginkan adalah 3 yakni kluster tingkat rendah, sedang, dan tinggi. Penentuan titik kluster ini dilakukan dengan mengambil nilai secara acak. Data awal akan diolah menggunakan rumus *Euclidian Distance* dengan *centroid* awal yang ditentukan secara acak. Berikut ini adalah *centroid* akhir dari hasil *k-means clustering* yang telah dilakukan.

Tabel 6. *Centroid Data*

Kluster	Rasio SD	Rasio SMP	Rasio SMU	IPM
(1)	(2)	(3)	(4)	(5)
0	2,95	2,86	2,83	4,31
1	2,48	2,41	2,78	4,21
2	2,72	2,72	2,85	4,22

Sumber : *kalbar.bps.go.id* (diolah)

Analisis Kluster

Pada penelitian ini, algoritma *k-means* digunakan untuk klasterisasi. Algoritma *K-means* merupakan algoritma yang bertujuan untuk mengelompokkan data menjadi beberapa kluster berdasarkan jarak terdekat. Analisis kluster dilakukan dengan jumlah kluster optimal sebanyak 3 kluster, diperoleh hasil sebagai berikut:

Tabel 7. Hasil Klasterisasi

Kluster	Anggota Kluster	Kabupaten/Kota
(1)	(2)	(3)
0	4	Ketapang, Kapuas Hulu, Melawi, Kayong Utara
1	2	Kota Singkawang, Kota Pontianak
2	8	Sambas, Bengkayang, Landak, Sanggau, Sintang, Kubu Raya, Mempawah, Sekadau

Sumber : *kalbar.bps.go.id* (diolah)

Berdasarkan tabel tersebut diperoleh bahwa terdapat empat daerah yang termasuk dalam klaster 0, yaitu Ketapang, Kapuas Hulu, Melawi, dan Kayong Utara. Ada dua daerah yang termasuk dalam klaster 1, yaitu Kota Singkawang dan Kota Pontianak. Ada delapan daerah yang termasuk dalam klaster 2, yaitu Sambas, Bengkayang, Landak, Sanggau, Sintang, Kubu Raya, Mempawah, dan Sekadau. Karakteristik tiap klaster dapat diketahui dengan membentuk *boxplot* berdasarkan rasio guru dengan murid di tingkat sd, smp, dan sma serta IPM.

Tabel 8. Tabel Karakteristik Tiap Klaster

Klaster	Rasio SD	Rasio SMP	Rasio SMU	IPM
(1)	(2)	(3)	(4)	(5)
0	Rendah	Rendah	Rendah	Rendah
1	Tinggi	Tinggi	Sedang	Tinggi
2	Sedang	Sedang	Tinggi	Sedang

Sumber : *kalbar.bps.go.id* (diolah)

Karakteristik klaster 0 berdasarkan rasio guru dengan murid pada tingkat sekolah dasar merupakan daerah dengan rasio berkategori rendah. Selanjutnya, karakteristik daerah yang termasuk klaster 1 merupakan daerah dengan rasio guru dengan murid berkategori tinggi. Terakhir, klaster 2 merupakan daerah dengan karakteristik rasio guru dengan murid berkategori sedang.

Karakteristik klaster 0 berdasarkan rasio guru dengan murid pada tingkat sekolah menengah pertama merupakan daerah dengan rasio guru dengan murid berkategori rendah. Selanjutnya, karakteristik daerah yang termasuk klaster 1 merupakan daerah dengan rasio berkategori sedang. Terakhir, klaster 2 merupakan daerah dengan karakteristik rasio guru dengan murid pada tingkat sekolah menengah pertama berkategori sedang. Karakteristik klaster 0 berdasarkan rasio guru dengan murid pada tingkat sekolah menengah umum berkategori rendah, Selanjutnya, karakteristik daerah yang termasuk klaster 1 merupakan daerah dengan rasio berkategori sedang. Terakhir, klaster 2 merupakan daerah dengan karakteristik rasio berkategori tinggi. Karakteristik terakhir berdasarkan indeks pembangunan manusia untuk klaster 0 adalah daerah dengan IPM berkategori rendah. Kemudian, klaster 1 berdasarkan IPM merupakan daerah dengan IPM berkategori Tinggi. Terakhir, klaster 2 berdasarkan IPM merupakan daerah dengan IPM berkategori sedang.

KESIMPULAN DAN SARAN

Berdasarkan penelitian ini, diperoleh tiga klaster untuk Kabupaten dan Kota di Provinsi Kalimantan Barat berdasarkan rasio antara guru dengan murid di tingkat sekolah dasar, sekolah menengah pertama, dan sekolah menengah umum serta indeks pembangunan manusia. Ada empat daerah yang termasuk dalam klaster 0, yaitu Ketapang, Kapuas Hulu, Melawi, dan Kayong Utara. Ada dua daerah yang termasuk dalam klaster 1, yaitu Kota Singkawang dan Kota Pontianak. Ada delapan daerah yang termasuk dalam klaster 2, yaitu Sambas, Bengkayang, Landak, Sanggau, Sintang, Kubu Raya, Mempawah, dan Sekadau.

Saran yang dapat peneliti berikan untuk daerah yang termasuk klaster 0 dengan karakteristik rasio SD, rasio SMP, rasio SMU, dan IPM rendah adalah dengan meningkatkan jumlah tenaga pengajar untuk tingkat SD, SMP, dan SMA. Dari segi IPM semakin tinggi nilai IPM maka mengindikasikan pembangunan yang baik pula sehingga diharapkan dengan adanya peningkatan jumlah tenaga pengajar dapat semakin meningkatkan IPM daerah pada klaster 0. Selanjutnya, untuk klaster 1 dengan karakteristik rasio SD tinggi, rasio SMP tinggi, rasio SMU sedang, dan IPM Tinggi dapat menjadi indikasi adanya kekurangan tenaga pengajar untuk tingkat SD dan SMP. Terakhir, saran untuk daerah yang termasuk ke dalam klaster 2 dengan karakteristik rasio SD sedang, rasio SMP sedang, rasio SMU

tinggi, dan IPM sedang adalah dengan meningkatkan jumlah tenaga pendidik untuk tingkat SD, SMP, dan SMU mengingat berada pada kategori sedang dan tinggi. Apabila tidak ada peningkatan jumlah tenaga pendidik dapat dikhawatirkan terjadi ketidakefektifan dalam proses belajar mengajar dan dapat berakibat pada penurunan nilai indeks pembangunan manusia.

DAFTAR PUSTAKA

- Agusta, Y. (2007). K-Means Penerapan, Permasalahan dan Metode Terkait, Jurnal Sistem dan Informatika, 3(1), pp,47-60.
- Dewa Ayu, K.P, & Dewa Ayu I.C.P. (2019) Analisis Perbandingan Metode Elbow dan Sillhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. Jurnal Matrix, Vol, 9, No.3
- Qori'atunnadyah, M. (2022). Pengelompokan Wilayah Berdasarkan Rasio Guru-Murid Pada Jenjang Pendidikan Menggunakan Algoritma K-Means. Jurnal Pengembangan Informatika.
- Arofah, I., & Rohimah, S. (2019). Analisis Jalur Untuk Pengaruh Angka Harapan Hidup, Harapan Lama Sekolah, Rata-Rata Lama Sekolah Terhadap Indeks Pembangunan Manusia Melalui Pengeluaran Riil Per Kapita Di Provinsi Nusa Tenggara Timur. Jurnal Saintika Unpam: Jurnal Sains Dan Matematika Unpam, 2(1), 76.
- Luthfi, K, & Taufiq,E. (2009). Algoritma Data Mining, Yogyakarta: Andi.
- M, Nishom. (2018). Implementasi Metode K-Means berbasis Chi-Square pada Sistem Pendukung Keputusan untuk Identifikasi Disparitas Kebutuhan Guru. J, Sist, Inf, Bisnis, vol, 8, no, 2, pp, 1–8.
- Febianto, N.I, & Palasara, N. (2019). Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018. Jurnal Sisfokom (Sistem Informasi dan Komputer), 8(2), 130-140. <https://doi.org/10.32736/sisfokom.v8i2.653>
- Rousseeuw, P. (1987). *Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis*".*Computational and Applied Mathematics*, 20: 53–65.
- Ramdhani, F. (2015). Pengelompokan Provinsi Di Indonesia Berdasarkan Karakteristik Kesejahteraan Rakyat Menggunakan Metode K-Means Cluster. GAUSSIANS Vol.4 No.4, 875-876.
- Sadewo, M.G., Windarto, A.P., dan Wanto, A. (2018). Penerapan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/ Mitigasi Bencana Alam Menurut Provinsi Dengan K-Means. Konferensi Nasional Teknologi Informasi dan Komputer, Vol 2 No 1.
- Sharma,G,dkk. (2016). *Analysis of K-Means clustering for Human Capital trends*. IEEE, 978-1-5090-5515-9/16/S31.00.