

## Pengelompokkan Wilayah Berdasarkan Rasio Guru-Murid Pada Jenjang Pendidikan Menggunakan Algoritma K-Means

Marita Qori'atunnadyah

Program Studi Informatika, Institut Teknologi dan Bisnis Widya Gama Lumajang

Corresponding Author: Marita Qori'atunnadyah (maritaqori@gmail.com)

### ARTICLE INFO

Date of entry:

1 Oktober 2022

Revision Date:

15 Oktober 2022

Date Received:

25 Oktober 2022

### ABSTRAK

Pendidikan memiliki peranan penting dalam kehidupan. Oleh karena itu, pemerintah perlu untuk mengupayakan pemerataan pendidikan. Salah satu indikator pemerataan pendidikan adalah rasio antara guru dan murid. Semakin kecil nilai rasio guru-murid berarti semakin efektif proses belajar mengajar yang dilakukan sehingga diharapkan mutu pendidikan semakin baik. Penelitian ini berfokus pada pengelompokkan wilayah berdasarkan rasio guru-murid pada jenjang pendidikan sekolah dasar dan menengah di Kabupaten Lumajang tahun 2021. Hasil yang didapatkan menunjukkan bahwa pengelompokkan wilayah terbagi menjadi 3 cluster. Cluster 1 merupakan cluster kecamatan dengan rasio guru-murid paling kecil yang memiliki anggota sebanyak 9 kecamatan. Kemudian cluster 2 merupakan cluster kecamatan dengan rasio guru-murid paling besar yang memiliki anggota sebanyak 2 kecamatan. Selanjutnya, cluster 3 merupakan cluster kecamatan dengan rasio guru-murid cukup yang memiliki anggota terbanyak yaitu 10 kecamatan. Berdasarkan hasil pengelompokkan tersebut, pemerintah Kabupaten Lumajang perlu memperhatikan kecamatan yang terdapat pada cluster 2 karena cluster tersebut memiliki rasio guru-murid yang paling besar, sehingga diharapkan kedepannya rasio guru-murid bisa menjadi lebih kecil.

Keywords: Guru, Murid, K-Means, Pengelompokkan, Rasio



### PENDAHULUAN

Pendidikan merupakan usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara (PMPK KEMDIKBUD, 2020). Dalam pembukaan Undang-Undang Dasar Negara Republik Indonesia tahun 1945 diamanatkan bahwa salah satu tujuan Pemerintah Negara Indonesia adalah untuk mencerdaskan kehidupan bangsa. Pendidikan menjadi landasan kuat yang diperlukan untuk meraih kemajuan bangsa di masa depan. Selain itu, pendidikan penting sebagai bekal dalam menghadapi era global dan persaingan ketat antarbangsa. Dengan demikian, pendidikan menjadi sangat penting untuk dipenuhi karena merupakan faktor bagi suatu bangsa untuk bisa bertahan dan unggul dalam

kompetisi global. Oleh karena itu, Pemerintah Negara Indonesia perlu mengupayakan pemerataan pendidikan. Pemerataan pendidikan ini berarti bahwa persamaan kesempatan bagi seluruh warga untuk mendapatkan pendidikan dan keadilan dalam memperoleh pendidikan yang sama dalam kehidupan masyarakat.

Beberapa faktor pendukung upaya pemerataan pendidikan antara lain sarana dan prasarana pendidikan, biaya pendidikan, kualitas tenaga pendidik, kurikulum, dan lain sebagainya. Salah satu faktor pendukung tersebut adalah guru. Guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan anak usia dini jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah (JDIH BPK RI, 2017). Oleh karena itu, peranan guru dalam pendidikan sangatlah penting. Pada proses pembelajaran, ada batasan jumlah tertentu di mana guru dapat membimbing peserta didiknya. Hal tersebut dapat menunjang efektivitas dan efisiensi selama proses pembelajaran. Perbandingan antara jumlah guru dan jumlah peserta didik disebut rasio guru-murid sesuai dengan yang terlampir di laporan Badan Pusat Statistik (BPS) Indonesia. Semakin kecil nilai rasio guru-murid maka semakin besar tingkat pengawasan dan perhatian guru terhadap murid yang berarti meningkatkan efektivitas dan efisiensi selama proses pembelajaran sehingga mutu pengajaran cenderung semakin baik. Sebagai contoh, rasio guru-murid untuk jenjang pendidikan SD yakni 1:15 yang berarti setiap 1 orang guru dapat mengajar maksimal 15 peserta didik di jenjang pendidikan SD. Pasal 17 Peraturan Pemerintah Nomor 74 Tahun 2008 tentang Guru menyebutkan bahwa pada jenjang SD, SMP, dan SMA idealnya satu guru bertanggung jawab terhadap 20 murid. Sedangkan, pada jenjang SMK idealnya satu guru bertanggung jawab pada 15 murid (SIMPUPH KEMENAG, 2020).

Kabupaten Lumajang merupakan salah satu daerah yang berada di wilayah bagian selatan Provinsi Jawa Timur. Wilayah ini terdiri dari 21 kecamatan, 198 desa dan 7 kelurahan. Pada tahun 2021, jumlah sekolah dasar dan menengah di Kabupaten Lumajang adalah 553 sekolah dasar (SD) dengan 4.642 guru dan 68.567 siswa, 137 sekolah menengah pertama (SMP) dengan 1.920 guru dan 28.687 siswa, serta 30 sekolah menengah atas (SMA) dengan 647 guru dan 11.318 siswa (Badan Pusat Statistik Kabupaten Lumajang, 2022). Upaya pemerataan pendidikan melalui jumlah guru di wilayah-wilayah perlu diperhatikan oleh Pemerintah Kabupaten Lumajang. Oleh karena itu, pengelompokan wilayah di Kabupaten Lumajang perlu untuk dilakukan agar pemerintah Kabupaten Lumajang dapat memperhatikan cluster wilayah yang membutuhkan perbaikan pendidikan khususnya berkaitan dengan rasio guru-murid. Cluster wilayah yang nantinya perlu diperhatikan adalah cluster dengan rasio guru-murid paling besar.

Salah satu metode yang digunakan untuk pengelompokan (clustering) dalam data mining adalah algoritma k-means. Algoritma ini termasuk dalam metode non-hierarchical clustering yang dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan, kemudian proses clustering dilakukan. Istilah k-means diusulkan oleh James McQueen pada tahun 1967. Algoritma ini pertama kali diperkenalkan oleh Stuart Lloyd pada tahun 1957 sebagai teknik modulasi kode pulsa. Algoritma clustering K-Means merupakan metode analisis cluster berbasis partisi (Vora & Oza, 2013). Algoritma k-means merupakan salah satu metode yang digunakan untuk mempartisi objek kedalam kelompok-kelompok berdasarkan kedekatan karakteristik, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain (Johnson & Wicherin, 2007). Penelitian yang menggunakan algoritma ini untuk pengelompokan wilayah yang berkaitan dengan pendidikan antara lain Pengelompokan Wilayah Berdasarkan Faktor Pendukung Pendidikan dengan Jumlah Sekolah dan Jumlah Guru menggunakan Algoritma K-Means (Dewi & Kamila, 2022), Pengelompokan Tingkat Pendidikan Berdasarkan Jumlah Sekolah Di Provinsi Jawa Barat Menggunakan Algoritma K-Means (Alfianti, 2022) dan Implementation of K-Means Clustering Method to Distribution of High School Teachers (Widiyaningtyas, Prabowo, & Pratama, 2017).

Berdasarkan uraian tersebut, penelitian ini berfokus pada Pengelompokan Wilayah Berdasarkan Rasio Guru-Murid Pada Jenjang Pendidikan Sekolah Dasar dan Menengah di Kabupaten Lumajang Tahun 2021 menggunakan Algoritma K-Means. Data yang digunakan dalam penelitian ini merupakan data olahan berupa rasio yang didapatkan dari jumlah guru dan murid pada jenjang pendidikan sekolah dasar dan menengah. Data didapatkan dari publikasi Badan Pusat Statistik Kabupaten Lumajang. Hasil penelitian ini diharapkan dapat menjadi masukan kepada Pemerintah Kabupaten Lumajang dalam merencanakan kebijakan dalam bidang pendidikan pada tahun berikutnya dengan memperhatikan cluster wilayah yang memiliki rasio guru-murid paling besar.

## METODE

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapatkan dari Badan Pusat Statistik Kabupaten Lumajang melalui publikasi “Kabupaten Lumajang Dalam Angka 2022”. Data yang digunakan terdiri dari 21 kecamatan meliputi data jumlah guru dan murid pada jenjang pendidikan Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), dan Sekolah Menengah Atas (SMA) (Badan Pusat Statistik Kabupaten Lumajang, 2022). Sebelum data digunakan dalam analisis menggunakan algoritma k-means, data diolah terlebih dahulu agar didapatkan nilai rasio guru dan murid.

Analisis cluster merupakan analisis yang bertujuan untuk mengelompokkan objek-objek berdasarkan kemiripan karakteristik yang dimilikinya. Tingkat kemiripan karakteristik yang dimaksud adalah objek-objek dalam suatu kelompok (cluster) sangat tinggi sedangkan tingkat kemiripan karakteristik objek antar cluster satu dengan yang lainnya rendah (Johnson & Wicherin, 2007). Terdapat 2 metode cluster yaitu hierarki dan non-hierarki. Dalam penelitian ini metode yang digunakan untuk mengelompokkan adalah non-hierarki.

Metode non-hierarki yang sering digunakan adalah algoritma k-means. Algoritma k-means merupakan salah satu metode analisis cluster non-hierarki yang dapat digunakan untuk mempartisi objek kedalam kelompok-kelompok berdasarkan kedekatan karakteristik, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain (Johnson & Wicherin, 2007). Kelebihan metode k-means adalah efisien untuk data yang besar. Kelemahan metode K-Means adalah jumlah/ banyaknya cluster dapat ditentukan di awal oleh peneliti.

Algoritma k-means mendefinisikan centroid dari sebuah cluster sebagai nilai rata-rata dari titik-titik di dalam cluster. Pertama, secara acak memilih  $k$  objek di  $D$ , yang masing-masing awalnya mewakili rata-rata atau pusat cluster. Untuk setiap objek yang tersisa, sebuah objek ditugaskan ke cluster yang paling mirip, berdasarkan jarak Euclidean antara objek dan rata-rata cluster. Algoritma k-means kemudian secara iteratif meningkatkan variasi dalam cluster. Untuk setiap cluster, rata-rata baru dihitung dengan menggunakan objek yang ditetapkan ke cluster pada iterasi sebelumnya. Semua objek kemudian dipindahkan menggunakan sarana yang diperbarui sebagai pusat cluster baru. Iterasi berlanjut hingga stabil, yaitu cluster yang terbentuk pada saat ini sama dengan yang terbentuk pada sebelumnya (Han, Kember, & Pei, 2012). Algoritma k-means dirangkum sebagai berikut.

1. Algoritma k-means untuk partisi, di mana setiap pusat cluster diwakili oleh nilai rata-rata objek dalam cluster.
2. Input:
  - $k$  : jumlah cluster,
  - $D$  : kumpulan data yang berisi  $n$  objek.
3. Output: satu set  $k$  cluster.
4. Algoritma:
  - a. Menentukan  $k$  (banyaknya cluster yang akan dibentuk) serta centroid awal di tiap cluster. Penentuan centroid awal dapat dilakukan secara acak.

- b. Menghitung jarak antara setiap objek dengan centroid awal, kemudian memasukkan objek-objek ke suatu cluster berdasarkan jarak terdekat dengan centroid yang bersesuaian. Umumnya perhitungan jarak dilakukan berdasarkan jarak euclidean sebagai berikut

$$dist = \sqrt{\sum_{j=1}^k \sum_{i=c_j} (x_i - z_j)^2},$$

dimana,  $c_j$  adalah cluster ke- $j$  dan  $z_j$  adalah centroid dari cluster  $c_j$  dan  $x_i$  adalah nilai objek (Thakare & Bagal, 2015).

- c. Menghitung kembali centroid dari cluster yang baru dibentuk.  
d. Mengulangi langkah (b) dan (c) sampai tidak ada lagi objek yang berpindah cluster.

## HASIL DAN PEMBAHASAN

Pada penelitian ini, analisis *cluster* dengan metode non-hierarki dilakukan menggunakan metode *k-means*. Dimana jumlah *cluster* yang ditentukan adalah sebanyak 3 *cluster*, sehingga pada analisis *cluster* dengan metode *k-means* digunakan  $k$  sebanyak 3 dengan data yang telah di *standardize* menggunakan *z-score*. Adapun hasil analisis metode *k-means* sebagai berikut.

**Tabel 1. Initial Cluster Centers Metode K-Means**

Cluster	Zscore(SD)	Zscore(SMP)	Zscore(SMA)
1	-0,32791	1,98978	0,20566
2	2,13143	0,17302	2,46327
3	-1,31165	-1,64373	-1,16854

Sumber: Data diolah (2022)

Tabel 1 memberikan informasi mengenai *centroid* awal atau inisial yang dipilih secara acak, sedangkan untuk banyak iterasi yang dibutuhkan untuk membentuk 3 *cluster* ditampilkan pada Tabel 2 berikut ini.

**Tabel 2. Jumlah Iterasi pada K-Means 3 Cluster**

Change in Cluster Centers	Iteration		
	1	2	3
1	1,340	0,133	0,000
2	1,054	0,000	0,000
3	1,199	0,118	0,000

Sumber: Data diolah (2022)

Berdasarkan hasil iterasi pada metode *k-means* dengan 3 *cluster* dapat diketahui bahwa untuk membentuk 3 *cluster* diperlukan 3 kali proses iterasi. Berikut disajikan pula nilai akhir pusat *cluster* pada Tabel 3.

**Tabel 3. Final Cluster Centers Metode K-Means**

Cluster	Zscore(SD)	Zscore(SMP)	Zscore(SMA)
1	0,27326	0,87954	0,18385
2	1,63956	0,17302	1,53078
3	-0,57385	-0,82619	-0,47162

Sumber: Data diolah (2022)

Tabel 3 diatas memberikan informasi mengenai *final cluster centers* yang digunakan untuk mengelompokkan observasi. Adapun jumlah observasi untuk setiap *cluster* yaitu sebagai berikut.

**Tabel 4. Jumlah Observasi tiap Cluster pada K-Means 3 Cluster**

Cluster	Jumlah
1	9
2	2
3	10
Valid	21
Missing	0

Sumber: Data diolah (2022)

Berdasarkan Tabel 4 dapat diketahui bahwa cluster ke-3 memiliki anggota terbanyak yaitu 10 kecamatan (Sumbersuko, Tekung, Kunir, Yosowilangun, Rowokangkung, Padang, Pasrujambe, Gucialit, Klakah, dan Ranuyoso), cluster ke-2 memiliki anggota paling sedikit yaitu 2 kecamatan (Lumajang dan Randuagung), dan cluster ke-1 memiliki anggota sebanyak 9 kecamatan (Tempursari, Pronojiwo, Candipuro, Pasirian, Tempeh, Jatiroto, Sukodono, Senduro, dan Kedungjajang). Cluster 1 merupakan cluster kecamatan dengan rasio guru-murid paling rendah, kemudian cluster 2 merupakan cluster kecamatan dengan rasio guru-murid paling tinggi, dan cluster 3 merupakan cluster kecamatan dengan rasio guru-murid cukup. Pada pendahuluan telah dijelaskan bahwa semakin kecil rasio guru-murid semakin baik. Sehingga pemerintah Kabupaten Lumajang perlu memperhatikan kecamatan yang terdapat pada cluster 2 agar rasio guru-murid menjadi lebih kecil dan proses belajar mengajar menjadi lebih efektif.

Terakhir, dilakukan pengujian ANOVA dan didapatkan hasil sebagai berikut.

**Tabel 5. ANOVA**

	Cluster		Error		F	Sig
	Mean Squae	df	Mean Squae	df		
Zscore(SD)	4,671	2	0,592	18	7,888	0,003
Zscore(SMP)	6,924	2	0,342	18	20,259	0,000
Zscore(SMA)	3,608	2	0,710	18	5,079	0,018

Sumber: Data diolah (2022)

Pada tabel ANOVA diatas diperoleh informasi bahwa *p-value* untuk semua variabel kurang dari taraf signifikan 0,05 yang artinya bahwa semua variabel (rasio guru-murid SD, SMP, SMA) berpengaruh signifikan terhadap pengelompokan.

## KESIMPULAN

Berdasarkan hasil dan analisis, didapatkan bahwa pengelompokan wilayah berdasarkan rasio guru-murid pada jenjang pendidikan sekolah dasar dan menengah di kabupaten lumajang tahun 2021 terbagi menjadi 3 cluster. Cluster 1 merupakan cluster kecamatan dengan rasio guru-murid paling kecil yang memiliki anggota sebanyak 9 kecamatan yaitu Tempursari, Pronojiwo, Candipuro, Pasirian, Tempeh, Jatiroto, Sukodono, Senduro, dan Kedungjajang. Kemudian cluster 2 merupakan cluster kecamatan dengan rasio guru-murid paling besar yang memiliki anggota sebanyak 2 kecamatan yaitu Lumajang dan Randuagung. Selanjutnya, cluster 3 merupakan cluster kecamatan dengan rasio guru-murid cukup yang memiliki anggota terbanyak yaitu 10 kecamatan yang terdiri dari Sumbersuko, Tekung, Kunir, Yosowilangun, Rowokangkung, Padang, Pasrujambe, Gucialit, Klakah, dan Ranuyoso. Berdasarkan hasil pengelompokan tersebut, pemerintah Kabupaten Lumajang perlu memperhatikan kecamatan yang terdapat pada cluster 2 karena cluster tersebut memiliki rasio guru-murid yang paling besar, sehingga diharapkan kedepannya rasio guru-murid bisa menjadi lebih kecil. Pengelompokan wilayah berdasarkan faktor pendidikan lainnya seperti jumlah sekolah, jenjang pendidikan Sekolah Menengah Kejuruan (SMK), Madrasah Ibtidaiyah (MI), Madrasah Tsanawiyah (MTs), dan Madrasah Aliyah (MA) juga dapat dilakukan untuk

penelitian selanjutnya. Metode pengelompokan lainnya seperti metode hierarki (simple linkage, complete linkage, dll) dan metode non-hierarki lainnya serta membandingkan beberapa metode tersebut sehingga didapatkan metode terbaik juga dapat dilakukan untuk penelitian selanjutnya.

## REFERENCES

- Alfianti, Z. I. (2022). Pengelompokan Tingkat Pendidikan Berdasarkan Jumlah Sekolah Di Provinsi Jawa Barat Menggunakan Algoritma K-Means. *JUSIKOM PRIMA (Jurnal Sistem Informasi dan Ilmu Komputer Prima)*, 5(2), 33-39.
- Badan Pusat Statistik Kabupaten Lumajang. (2022). *Kabupaten Lumajang Dalam Angka 2022*. Lumajang: Badan Pusat Statistik Kabupaten Lumajang.
- Dewi, E. T., & Kamila, I. (2022). Pengelompokan Wilayah Berdasarkan Faktor Pendukung Pendidikan dengan Jumlah Sekolah dan Jumlah Guru menggunakan Algoritma K-Means. *INTERVAL: Jurnal Ilmiah Matematika*, 2(1), 1-12.
- Han, J., Kember, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed.). USA: Morgan Kaufmann.
- JDIH BPK RI. (2017). *DATABASE PERATURAN*. Retrieved November 28, 2022, from <https://peraturan.bpk.go.id/Home/Details/40266/uu-no-14-tahun-2005>
- Johnson, R., & Wicherin, D. (2007). *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Prentice Hall.
- PMPK KEMDIKBUD. (2020). *Direktorat Pendidikan Masyarakat dan Pendidikan Khusus, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, Republik Indonesia*. Retrieved November 26, 2022, from <https://pmpk.kemdikbud.go.id/>
- SIMPUH KEMENAG. (2020). *PP 74 Tahun 2008*. Retrieved November 28, 2022, from [https://simpuh.kemenag.go.id/regulasi/pp\\_74\\_08.pdf](https://simpuh.kemenag.go.id/regulasi/pp_74_08.pdf)
- Thakare, Y. S., & Bagal, S. B. (2015). Performance Evaluation of K-means Clustering Algorithm with Various Distance Metrics. *International Journal of Computer Applications*, 110(11), 13-16.
- Vora, P., & Oza, B. (2013). A Survey on K-mean Clustering and Particle Swarm Optimization. *International Journal of Science and Modern Engineering*, 1(3), 24-26.
- Widiyaningtyas, T., Prabowo, M. I., & Pratama, M. A. (2017). Implementation of K-Means Clustering Method to Distribution of High School Teachers. In: *2017 4th nternational Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, (pp. 1-6). Yogyakarta.