



DASHBOARD LINGKUNGAN HIDUP UNTUK ANALISIS DIARE MENGUNAKAN METODE K-MEANS CLUSTERING

Sitti Sahara¹, Saeful Amri², Ariska Fitriyana Ningrum³, Dannu Purwanto⁴

1234Fakultas Sains dan Teknologi Pertanian Universitas Muhammadiyah Semarang, Indonesia

DOI: <https://doi.org/10.26714/jodi>

Info Artikel

Sejarah Artikel:

Disubmit 27 Oktober 2023

Direvisi 19 Juli 2024

Disetujui 30 Juni 2024

Keywords:

Lingkungan; diare; k-means.

Abstrak

Diare adalah penyakit umum dengan penyebab yang beragam, termasuk virus, bakteri, dan faktor-faktor lainnya. Faktor-faktor lingkungan, gizi yang buruk, dan kurangnya pengetahuan masyarakat berperan penting dalam tingginya kasus diare, terutama pada anak-anak di bawah lima tahun, di Indonesia. Analisis cluster digunakan untuk mengelompokkan daerah berdasarkan kasus diare dan membantu perencanaan penanggulangan. Penelitian ini menggunakan data BPS 2021 dari 34 provinsi di Indonesia dan berfokus pada faktor penyebab diare. Penelitian ini bertujuan untuk memahami faktor-faktor yang berkontribusi pada kasus diare, dengan harapan dapat merumuskan strategi penanggulangan yang lebih efektif

Abstract

Diarrhea is a common disease with diverse causes, including viruses, bacteria, and other factors. Environmental factors, poor nutrition, and lack of public knowledge play an important role in the high incidence of diarrhea, especially in children under five years old, in Indonesia. Cluster analysis is used to group regions based on diarrhea cases and help with prevention planning. This study used 2021 BPS data from 34 provinces in Indonesia and focused on the causal factors of diarrhea. This study aims to understand the factors that contribute to diarrhea cases, with the hope of formulating more effective prevention strategies.

PENDAHULUAN

Diare, atau yang berasal dari kata Yunani "Diarroi," adalah penyakit yang umumnya dapat menyerang individu dari berbagai kelompok usia, termasuk remaja, dewasa, dan balita. Gejalanya meliputi peningkatan volume, keenceran, dan frekuensi buang air besar yang melebihi ambang batas tertentu, yaitu lebih dari 3 kali sehari pada orang dewasa dan lebih dari 4 kali sehari pada bayi neonatus, seringkali dengan lendir darah. Penyebab diare bervariasi, termasuk virus seperti Rotavirus (40-60%), bakteri seperti *Escherichia Coli* (20-30%), *Shigella Sp* (1-2%), dan parasit seperti *Entamoeba Hystolitica* (<1%). Faktor lain yang dapat memicu diare mencakup keracunan makanan, infeksi bakteri, penggunaan pemanis buatan, dan konsumsi fruktosa. Diare juga dipengaruhi oleh faktor-faktor seperti kurangnya higiene dan sanitasi, malnutrisi, lingkungan yang padat, serta keterbatasan sumber daya medis. Organisasi Kesehatan Dunia (WHO) mencatat bahwa diare adalah penyebab kematian kedua yang paling umum pada anak-anak di bawah usia lima tahun. Di Indonesia, prevalensi diare merupakan masalah kesehatan masyarakat dengan tingginya jumlah kasus. Beberapa faktor penerimaan (faktor pejamu) dapat meningkatkan insiden diare, termasuk faktor lingkungan yang tidak sehat yang memungkinkan penyebaran kuman penyebab diare dan perilaku manusia yang tidak sehat. Persediaan air bersih yang terbatas juga dapat menjadi faktor yang mempermudah penyebaran penyakit di masyarakat. Akses yang terbatas terhadap air bersih dan sanitasi merupakan pelanggaran hak asasi manusia dan menjadi dasar bagi kesehatan yang baik. Setiap tahun, lebih dari 3,5 juta orang di seluruh dunia meninggal akibat penyakit menular yang terkait dengan air dan kontaminasi tinja, termasuk diare. Gizi yang buruk juga menjadi penyebab utama diare pada balita. Setiap tahun, terdapat sekitar 1,7 miliar kasus penyakit diare pada anak-anak. Anak-anak yang mengalami kekurangan gizi atau memiliki sistem kekebalan tubuh yang terganggu, serta individu yang hidup dengan HIV, berisiko tinggi mengalami diare yang berpotensi mengancam nyawa. Sementara di masa lalu, dehidrasi yang berat sering menjadi komplikasi utama diare, saat ini, penyebab diare yang berhubungan dengan kematian lebih sering adalah infeksi bakteri septik. Kurangnya pengetahuan dan kesadaran masyarakat merupakan faktor penting dalam meningkatnya kasus diare. Upaya penanggulangan diare perlu dimulai di tingkat daerah dengan kerja sama antara pihak berwenang dan masyarakat. Analisis cluster merupakan metode yang dapat digunakan untuk mengelompokkan daerah berdasarkan kasus diare dan permasalahan terkaitnya, yang dapat memberikan panduan bagi strategi penanggulangan yang lebih efektif. Oleh karena itu, penting untuk mengidentifikasi penyebab dan jumlah penderita diare di setiap daerah sebagai langkah yang mendesak dalam penanganannya.

METODE

Menurut Widodo (2013:9) Clustering atau klasifikasi adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Cluster adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam cluster yang sama dan disimilar terhadap objek-objek yang berbeda cluster. Analisis cluster bermanfaat dalam menyimpulkan suatu data yang kompleks dengan cara mengelompokkan objek-objek yang memiliki kemiripan karakteristik. Proses clustering merupakan proses yang dilakukan dengan dua metode, yaitu dengan metode hierarki dan non hierarki:

1. Metode hierarki (hierarchical method) Metode hierarki (hierarchical method) merupakan metode yang dilakukan secara bertahap. Pada metode ini akan membentuk tahapan tertentu seperti pada struktur pohon serta dapat di hasilkan

bentuk dendogram. Dendogram merupakan representasi visual dari tahap-tahap proses analisis cluster yang terbentuk dan nilai koefisien jarak pada setiap tahap.

2. Metode Non Hierarki (Non Hierarchical Method) Metode non hierarki disebut juga dengan metode k-means. Metode ini tidak sama halnya dengan metode hierarki, karena pada metode non hierarki dimulai dengan menentukan terlebih dahulu sejumlah cluster awal yang diinginkan, K-Means Clustering merupakan metode pengklasteran secara partitioning yang memisahkan data ke dalam kluster yang berbeda. Dengan partitioning secara iteratif, K-means Clustering mampu meminimalkan rata-rata jarak setiap data ke kluster terdekat. Langkah-langkah untuk KMeans Clustering adalah sebagai berikut (Prasetyo, 2014):

- 1) Tentukan k sebagai jumlah cluster yang akan dibentuk.
- 2) Tentukan k centroid (titik pusat cluster) awal secara random/acak. Penentuan pada centroid awal dilakukan secara acak atau random dari objek yang tersedia sebanyak k-cluster.
- 3) Hitung jarak setiap objek ke masing-masing centroid dari masing-masing cluster. Untuk menghitung jarak antara objek dengan centroid dapat menggunakan Euclidean Distance, dengan rumus sebagai berikut:

$$(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2}$$

- 4) Alokasikan masing-masing objek ke dalam centroid yang paling dekat.
- 5) Lakukan iterasi, kemudian tentukan posisi centroid baru dengan menggunakan persamaan untuk menghitung centroid cluster ke-i berikutnya, digunakan rumus sebagai berikut:

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

Dimana :

v = centroid pada cluster

x_i = jumlah objek cluster ke-i

n = banyaknya objek yang menjadi anggota cluster

- 6) Ulangi langkah 3 jika posisi centroid baru tidak sama.

Data yang digunakan pada penelitian ini adalah data sekunder yang diambil dari Bada Pusat Statistik (BPS) pada 34 Provinsi yang ada di Indonesia tahun 2021 yang disediakan oleh PT.Maribelajar Indonesia Cerdas sebagai bahan Capstone Project. Jenis data yang digunakan yaitu data kuantitatif dan variabel penelitian yang digunakan terdiri dari Perkiraan kasus diare, kasus diare dilayani, dan persentase kasus diare dilayani. Untuk analisis dalam penelitian ini yaitu menggunakan software IBM SPSS 22 dan Microsoft Excel. Adapun prosedur penelitian yang dilakukan yaitu sebagai berikut:

1. Input data
2. Pengujian data
3. Uji asumsi multikolinearitas
4. Standarisasi data
5. Uji K-Means Cluster
6. Interpretasi hasil
7. Selesai.

HASIL DAN PEMBAHASAN
Uji Multikolinearitas

Table 1 Hasil Uji Multikolinearitas

Model	Collinearity Statistics	
	Tolerance	VIF
1 Perkiraan_kasus_diare	.872	1.146
Kaus_diare_dilayani	.804	1.244
Presentase_kasus_diare_dilayani	.843	1.186

Tabel di atas merupakan hasil pengujian dari multikolinieritas. Didapatkan bahwa variabel Perkiraan Kasus Diare, Kasus Diare Dilayani dan Presentase Kasus Diare Dilayani memiliki nilai tolerance lebih besar dari 0,1 serta nilai VIF nya kurang dari 10. Maka dapat disimpulkan bahwa semua variabel tidak terjadi multikolinieritas. Karena data tidak terjadi multikolinearitas, maka Transformasi data ke dalam bentuk logaritma natural tidak perlu dilakukan, dan dilanjutkan ke analisis K-Means. Analisis Cluster K-Means Pada penelitian ini akan dilakukan pengujian analisis cluster dengan metode K-Means untuk clustering Provinsi di Indonesia berdasarkan tingkat perkiraan kasus penyakit diare menjadi 3 cluster, yaitu tingkat kasus persebaran tinggi, kasus persebaran sedang, dan kasus persebaran rendah.

Table 2 Iteration History

Iteration	Change in ClusterCenters		
	1	2	3
1	.327	.000	1.556
2	.000	.000	.000

a. Convergence achieved due to noor small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 2. The minimum distance between initial centers is 5.262.

Tabel di atas merupakan proses iterasi dalam pengelompokan cluster dari tabel initial dan menghasilkan proses iterasi sebanyak 2 kali. Pada iterasi 1 terjadi centeroid yang tidak signifikan dan pada iterasi 2 terjadi centeroid yang signifikan. Jadi, semua cluster sudah terbentuk dan iterasi berhenti pada iterasi 2 dengan jarak minimum 5,262.

Table 3 Final Cluster Center

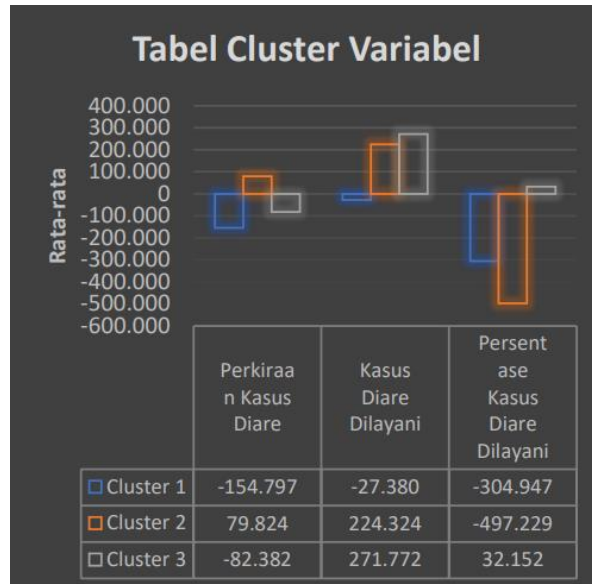
	Cluster		
	1	2	3
Perkiraan_kasus_diare	-.09888	4.66352	-.56573
Kaus_diare_dilayani	-.29077	1.88220	2.28032
Presentase_kasus_diare_dilayani	-.20549	-.33506	2.16657

Pada tabel di atas adalah hasil dari proses akhir dalam clustering yang membentuk cluster sebanyak 3 untuk masing-masing variabel. Variabel pada tabel Final Cluster Centers merupakan hasil untuk nilai standarisasi. Angka-angka yang terbentuk di atas mempunyai arti. Angka negatif memiliki arti bahwa data berada di bawah rata-rata total dan angka positif memiliki arti bahwa data berada di atas rata-rata total. Untuk mengetahui pengaruh pada variabel di atas dilakukan perhitungan dengan rumus:

$$X = \mu + z. \sigma$$

Keterangan:

X = rata-rata data atau sampel
 μ = rata-rata populasi
 σ = standar deviasi
 z = nilai standarisasi.



Gambar 2 Rata-rata Cluster Pada Masing-masing Variabel

Berdasarkan perhitungan yang dihasilkan dari masing-masing variabel, didapatkan karakteristik pada masing-masing cluster. Berikut adalah penjelasan untuk setiap cluster:

1. Cluster 1
 Dari gambar grafik diatas dapat diketahui bahwa karakteristik yang terdapat pada cluster 1 berisi variabel perkiraan kasus diare, kasus diare dilayani dan persentase kasus diare dilayani, semua variable masing-masing berada di bawah rata-rata populasi.
2. Cluster 2
 Dari gambar grafik diatas dapat diketahui bahwa karakteristik yang terdapat pada cluster 2 berisi variabel perkiraan kasus diare dan kasus diare dilayani berada di atas rata-rata populasi, sedangkan variabel persentase kasus diare dilayani berada di bawah ratarata populasi.
3. Cluster 3
 Dari gambar grafik diatas dapat diketahui bahwa karakteristik yang terdapat pada cluster 3 berisi variabel perkiraan kasus diare yang berada di bawah rata-rata populasi, sedangkan variabel kasus diare dilayani dan persentase kasus diare dilayani berada di atas rata-rata populasi.

Jumlah Anggota Masing-Masing Cluster

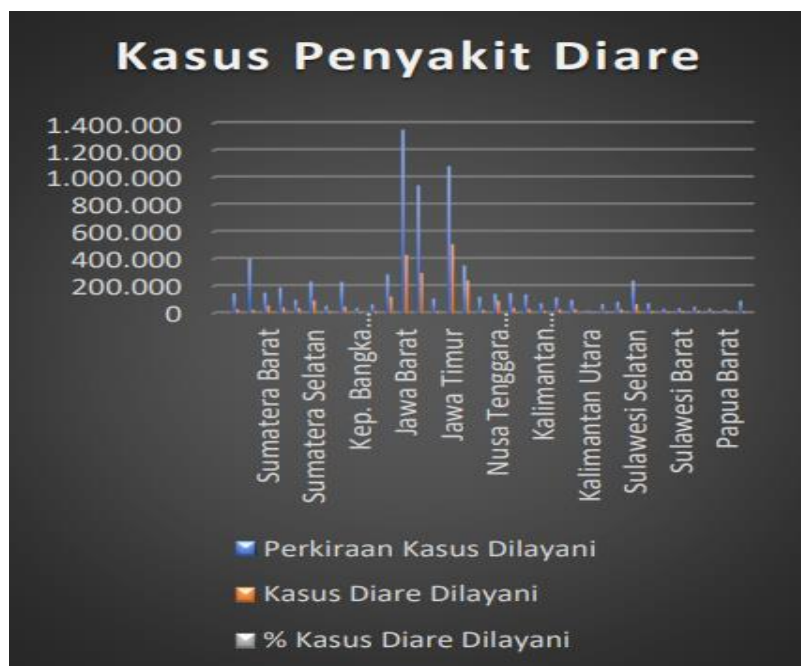
Table 4 Number of cases in Each Cluster

Number of Cases in each Cluster	
Cluster	1
	2
	3
Valid	34.000
Missing	.000

Berdasarkan tabel di atas, tidak ada data yang missing atau hilang, dan pada cluster 1 terdapat jumlah anggota sebanyak 30 provinsi, cluster 2 sebanyak 1 provinsi, dan cluster 3 sebanyak 3 provinsi.

Profilling Komposisi Cluster

Setelah mengetahui jumlah anggota untuk masing-masing cluster, didapatkan tabel cluster membership yang didapatkan dari Hasil Ouput SPSS Analisis K-Means Cluster pada kelompok cluster. Berikut adalah tabel berisi Provinsi pada masingmasing cluster dan gambar berisi Provinsi yang memiliki kasus tertinggi pada setiap cluster.



Gambar 4.2 Kasus Penyakit Diare Di Setiap Provinsi

Berdasarkan gambar diatas, dapat diketahui bahwa kasus penyakit diare tertinggi terdapat pada Provinsi Jawa Barat yaitu perkiraan kasus diare dilayani sebanyak 1,352,788 kasus, dengan kasus diare dilayani sebanyak 430,520 kasus, dan untuk presentase kasus diare dilayani yaitu sebesar 31,8%, sedangkan kasus penyakit diare terendah terdapat pada Provinsi Aceh dengan perkiraan kasus diare dilayani yaitu sebanyak 147,396 kasus, kasus diare dilayani yaitu sebanyak 29,866, dan kasus

persentase kasus diare dilayani yaitu sebesar 20,3% kasus. Selanjutnya karakteristik kelompok setiap Provinsi dapat dilihat pada table 5 dibawah ini.

Table 5 Karakteristik Kelompok Provinsi KMeans

Cluster	Karakteristik	Anggota Cluster	Jumlah
1	Kasus Penyakit Diare Rendah	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua	30
2	Kasus Penyakit Diare Sedang	Jawa Tengah	1
3	Kasus Penyakit Diare Tinggi	Jawa Barat, Jawa Timur, Nusa Tenggara Barat	3

Provinsi yang masuk dalam satu cluster pada variabel kasus penyakit diare merupakan wilayah-wilayah yang memiliki kesamaan dan kemiripan karakteristik berdasarkan data yang ada. Dimana kelompok 1 terdiri dari Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua, dimana masuk dalam presentase tingkat kasus penyakit diare rendah pada Provinsi tersebut.

Pada kelompok 2 yang terdiri dari Provinsi Jawa Tengah dimana masuk dalam presentase tingkat kasus penyakit diare sedang. Dan pada kelompok 3 yang terdiri dari Provinsi Jawa Barat, Jawa Timur dan Nusa Tenggara Barat, dimana masuk dalam presentase tingkat kasus penyakit diare tinggi. Pada cluster ini menjadi prioritas dalam pencegahan bertambahnya kasus diare.

Pembahasan

Diare adalah gangguan pencernaan yang ditandai dengan buang air besar terus menerus. Fases yang keluar saat BAB biasanya lembek atau cair. Diare adalah buang air besar dengan bertambahnya frekuensi yang lebih dari biasanya 3 kali sehari atau lebih dengan konsistensi cair. Penyakit diare merupakan salah satu masalah kesehatan masyarakat di Indonesia. Hal ini dikarenakan masih tingginya angka kesakitan diare yang menimbulkan kematian terutama pada balita. Penyebab diare pada balita tidak dapat dilepaskan dari kebiasaan hidup sehat dari setiap keluarga dan sanitasi lingkungan yang tidak sehat. Faktor penyebab terjadinya penyakit diare pada balita yaitu berhubungan dengan faktor lingkungan. Faktor lingkungan tersebut meliputi sarana air bersih yang berhubungan langsung dengan kebutuhan manusia sehari-hari seperti untuk minum, mandi, mencuci, dan lain sebagainya. Jika air terkontaminasi dengan kotoran, seperti kebiasaan perilaku tidak mencuci tangan dengan sabun setelah BAB dari jamban atau tempat pembuangan tinja, pembuangan limbah pabrik ataupun limbah rumah tangga tidak pada tempatnya akibatnya akan mencemari lingkungan. Hal tersebut akan berdampak buruk bagi kesehatan masyarakat. Dengan faktor lingkungan yang buruk tersebut dapat menyebabkan seorang balita terkena diare.

KESIMPULAN

Berdasarkan penelitian yang dilakukan menggunakan analisis K-Means Clustering, dapat diambil kesimpulan bahwa penanganan kasus penyakit diare pada beberapa provinsi belum terlaksana dengan baik, khususnya pada kota-kota besar yang termasuk pada kasus penyakit diare tertinggi, faktor tingginya kasus diare suatu daerah yaitu salah satunya faktor lingkungan yang meliputi penggunaan air bersih, penggunaan jamban (pembuangan kotoran/tinja), kebersihan rumah, dan pembuangan limbah. Karena penggunaan sanitasi lingkungan tidak dapat dilepaskan dari aktivitas kehidupan manusia, jika keadaan sanitasi lingkungan kotor atau tidak sehat, maka akan menyebabkan terjadinya masalah besar. Salah satunya yaitu dapat menyebabkan mewabahnya penyakit diare khususnya pada balita.

DAFTAR PUSTAKA

- Iryanto, A.G, dkk. 2021. "Literatur Review: Faktor Resiko Kejadian Diare Pada Balita Di Indonesia": Jurnal Kesehatan Lingkungan Vol 11. Universitas Diponegoro Semarang Indonesia.
- Miswan, dkk. 2018. "Hubungan Sanitasi Lingkungan Dengan Penyakit Diare Pada Masyarakat Di Desa Tumpapa Indah Kecamatan Balinggi Kabupaten Parigi Moutong Provinsi Sulawesi Tengah": UNM Environmental Journals Volume 1. Universitas Negeri Makassar.
- Nurrohmah, M. Femila, L. 2020. "Hubungan Antara Sanitasi Lingkungan Dengan Kejadian Diare Pada Balita", <https://osf.io/cqkqm>, diakses pada 25 July 2023 pukul 16.11.
- Nugraha, P, dkk. 2022. "Faktor-Faktor Yang Berhubungan Dengan Kejadian Diare Pada Balita Di Kelurahan Kapuas Kanan Hulu Kecamatan Sintang": Jurnal Kesehatan Masyarakat Vol 1. Sekolah Tinggi Ilmu Kesehatan Kapuas Raya Sintang.
- Qisti, D.A, dkk. 2021. "Analisis Aspek Lingkungan Dan Perilaku Terhadap Kejadian Diare Pada Balita Di Tanah Sareal": Jurnal Inovasi Penelitian Vol 2. Sekolah Tinggi Ilmu Kesehatan Indonesia Maju.
- Sari, Desy Noor Permata. 2021. "Analisis Cluster Dengan Metode K-Means Pada Persebaran Kasus Covid-19 Berdasarkan Provinsi Di Indonesia", <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/45035>, diakses pada 26 July 2023 pukul 22.33.