

EKSPLORASI POLA PENYAKIT MENULAR DI NUSA TENGGARA TIMUR MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Nabila Nur Rosyada^{1*}, Muhammad Bahtiar Afandi², Mahmudah³

Departemen Kesehatan Masyarakat, Fakultas Kesehatan Masyarakat, Universitas Airlangga, Surabaya, Indonesia^{1,2}, Departemen Epidemiologi, Biostatistika, Kependudukan dan Promosi Kesehatan Fakultas Kesehatan Masyarakat, Universitas Airlangga, Surabaya, Indonesia³

*Corresponding Author : nabila.rosyada13@gmail.com

ABSTRAK

Penyakit menular memberikan beban yang jauh lebih berat pada daerah kepulauan seperti Provinsi Nusa Tenggara Timur di Indonesia, di mana kerentanan lingkungan dan ketimpangan layanan kesehatan semakin memperparah variasi spasial pola kejadian kasus. Penelitian ini bertujuan mengidentifikasi dan mengkarakterisasi kelompok kabupaten untuk delapan penyakit menular utama menggunakan metode klasterisasi K-means. Data sekunder mengenai malaria, tuberkulosis, pneumonia, kusta, diare, demam berdarah dengue, AIDS, serta infeksi menular seksual dari 22 kabupaten diambil dari Badan Pusat Statistik dan dianalisis menggunakan R Studio (v. 2025.05.1); variabel divalidasi melalui standarisasi z-score setelah pengujian KMO (0,61) dan VIF, dengan jumlah klaster optimal $k=3$ ditentukan melalui metode elbow dan silhouette. Terdapat tiga klaster yang terbentuk: Klaster 1 (4 kabupaten di wilayah timur) didominasi kasus malaria (rata-rata 1.701) dan diare (2.789), Klaster 2 (Kota Kupang) dengan AIDS (2.577), IMS (604), serta diare (4.476), sedangkan Klaster 3 (17 kabupaten) menunjukkan beban penyakit yang sedang dan lebih merata (total rata-rata 1.675). Visualisasi radar plot dan principal component mengonfirmasi profil yang berbeda antar klaster. Klasterisasi K-means ini mengungkap pola spasial yang dapat ditindaklanjuti untuk mendukung intervensi yang tepat sasaran di daerah dengan sumber daya terbatas.

Kata kunci : epidemiologi spasial, klasterisasi K-means, Nusa Tenggara Timur, penyakit menular, pola penyakit

ABSTRACT

Infectious diseases impose a disproportionate burden on archipelagic regions like East Nusa Tenggara Province, Indonesia, where environmental vulnerabilities and healthcare disparities exacerbate spatial heterogeneity in incidence patterns. This study aimed to identify and characterize district-level clusters of eight major communicable diseases using K-means clustering. Secondary data on malaria, tuberculosis, pneumonia, leprosy, diarrhea, dengue, AIDS, and sexually transmitted infections from 22 districts were sourced from Statistics Indonesia and analyzed in R Studio (v. 2025.05.1); variables underwent z-score standardization following KMO (0,61) and VIF validation, with optimal $k=3$ determined via elbow and silhouette methods. Three clusters emerged: Cluster 1 ($n=4$ eastern districts) dominated by malaria (mean=1,701) and diarrhea (2,789), Cluster 2 (Kota Kupang) by AIDS (2,577), IMS (604), and diarrhea (4,476), and Cluster 3 ($n=17$) with moderate equilibrated burdens (total mean=1,675). Radar and PC plots confirmed distinct profiles. K-means clustering reveals actionable spatial patterns to guide targeted interventions in resource-constrained settings.

Keywords : K-means clustering, infectious diseases, spatial epidemiology, Nusa Tenggara Timur, disease patterns

PENDAHULUAN

Penyakit menular tetap menjadi masalah utama dalam kesehatan masyarakat dunia, khususnya di negara-negara berpenghasilan rendah dan menengah yang menghadapi kerentanan lingkungan, ketimpangan sosial-ekonomi, serta keterbatasan sistem kesehatan. Meskipun ada kemajuan signifikan dalam upaya pengendaliannya, penyakit infeksi ini masih

menyumbang porsi besar angka kesakitan dan kematian di seluruh dunia, sekaligus memicu dampak sosial serta beban ekonomi yang berat (World Health Organization, 2022; Murray et al., 2020). Tantangan yang terus berlanjut ini menekankan pentingnya pendekatan analisis yang lebih canggih, yang tidak hanya mengandalkan data agregat tetapi mampu menangkap variasi spasial dalam pola penyebaran penyakit (Frenk et al., 2017). Di Indonesia, beban penyakit menular masih belum merata antar wilayah seiring berlangsungnya transisi epidemiologi. Penyakit seperti malaria, tuberkulosis, pneumonia, diare, dan demam berdarah dengue cenderung lebih parah menyerang provinsi-provinsi dengan infrastruktur terbatas dan akses layanan kesehatan yang minim (Ministry of Health of Indonesia, 2023). Provinsi Nusa Tenggara Timur menjadi salah satu contoh wilayah yang sangat rentan akibat karakteristik geografisnya yang berbentuk kepulauan, iklim kering, ketersediaan air bersih terbatas, serta distribusi layanan kesehatan yang timpang. Kondisi-kondisi tersebut memicu penularan berkelanjutan berbagai penyakit menular, termasuk malaria, tuberkulosis, pneumonia, diare, kusta, infeksi menular seksual, dan HIV/AIDS, dengan variasi yang mencolok antar kabupaten dan kota (Statistics Indonesia, 2023; Suryani et al., 2020; Titaley et al., 2022).

Selain faktor geografis dan infrastruktur, perilaku kesehatan masyarakat juga memengaruhi persebaran penyakit menular. Praktik sanitasi yang buruk, kurangnya kesadaran penggunaan air bersih, serta kepatuhan rendah terhadap program imunisasi berkontribusi terhadap tingginya kejadian infeksi di tingkat komunitas (Putri et al., 2021; Titaley et al., 2022). Kondisi ini diperparah oleh rendahnya kapasitas sistem kesehatan daerah untuk melakukan surveilans, deteksi dini, dan respons cepat terhadap wabah lokal, sehingga intervensi kesehatan masyarakat sering terlambat dan kurang efektif (Ministry of Health of Indonesia, 2023; Suryani et al., 2020). Perkembangan teknologi analisis data menawarkan solusi potensial dalam memetakan risiko penyakit menular secara lebih tepat. Metode unsupervised machine learning, khususnya algoritma K-Means, mampu mengidentifikasi pola kluster penyakit berdasarkan karakteristik epidemiologis dan demografis wilayah. Pendekatan ini telah berhasil diterapkan dalam berbagai studi untuk mendukung pengambilan keputusan yang lebih terfokus, seperti alokasi sumber daya kesehatan dan perencanaan intervensi berbasis bukti (Jain, 2016; Han et al., 2022).

Dengan memahami kluster penyakit di tingkat kabupaten, intervensi dapat disesuaikan dengan kondisi lokal, sehingga meningkatkan efektivitas pengendalian penyakit menular. Selain itu, analisis spasial yang dihasilkan dari klusterisasi juga membantu memvisualisasikan kesenjangan layanan kesehatan dan faktor risiko lingkungan yang berbeda antar wilayah. Dengan informasi ini, pemangku kebijakan dapat merancang strategi pencegahan yang lebih adaptif, misalnya melalui distribusi fasilitas kesehatan bergerak, program edukasi komunitas, dan kampanye kesehatan berbasis lokasi (Frenk et al., 2017; Putri et al., 2021). Hal ini menegaskan pentingnya pemanfaatan data dan teknologi dalam mengatasi tantangan penyakit menular yang kompleks di wilayah kepulauan seperti Nusa Tenggara Timur.

Dengan latar belakang tersebut, penelitian ini bertujuan untuk mengidentifikasi dan mengkarakterisasi pola penyakit menular pada tingkat kabupaten di Provinsi Nusa Tenggara Timur menggunakan klusterisasi K-Means. Hasil klusterisasi yang disajikan menjadi dasar untuk menginterpretasikan profil penyakit regional, yang kemudian dibahas lebih lanjut terkait kondisi lingkungan, kapasitas sistem kesehatan, serta implikasi untuk pengendalian penyakit menular.

METODE

Penelitian ini menggunakan rancangan observasional kuantitatif dengan pendekatan *ecological cross-sectional* yang memanfaatkan data sekunder untuk menganalisis pola penyakit menular pada tingkat kabupaten di Provinsi Nusa Tenggara Timur, Indonesia. Seluruh

22 kabupaten dan kota di provinsi tersebut dijadikan unit analisis melalui total sampling, sehingga semua wilayah administratif tercakup sepenuhnya. Data yang digunakan bersumber dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur tahun 2024, mencakup jumlah kasus agregat penyakit menular seperti malaria, tuberkulosis (TBC), pneumonia, diare, demam berdarah dengue (DBD), kusta, HIV/AIDS, serta infeksi menular seksual (IMS), yang dikumpulkan melalui mekanisme pelaporan standar dan verifikasi antarlembaga, menjamin validitas serta reliabilitas data sekunder.

Pengolahan data dilakukan menggunakan perangkat lunak statistik R versi 2025.09.2+418, dimulai dari pembersihan data, pengecekan kelengkapan, hingga seleksi variabel yang relevan. Standarisasi z-score diterapkan agar setiap variabel memiliki skala sebanding, sesuai praktik terbaik untuk metode klasterisasi berbasis jarak (Jain, 2016; Han et al., 2022). Sebelum dilakukan klasterisasi, kelayakan dataset diuji menggunakan ukuran Kaiser-Meyer-Olkin (KMO) dan analisis korelasi antarvariabel untuk memastikan tidak ada masalah multikolinearitas sehingga struktur data cocok untuk analisis multivariat. Langkah ini bertujuan agar klaster yang dihasilkan dapat merepresentasikan pola penyakit yang sebenarnya dengan akurat.

Analisis klasterisasi dilakukan menggunakan algoritma K-Means, metode unsupervised machine learning yang banyak diterapkan dalam epidemiologi dan penelitian kesehatan masyarakat (MacQueen, 1967; Hartigan & Wong, 1979; Jain, 2016). Jumlah klaster optimal ditentukan melalui kombinasi metode Elbow dan Silhouette coefficient untuk mengevaluasi kekompakan dan pemisahan antar klaster (Rousseeuw, 1987). Setelah pengelompokan, profil klaster dianalisis dengan menghitung rata-rata tiap variabel penyakit untuk mengidentifikasi pola dominan di setiap klaster (Putri et al., 2021; Prasetyo et al., 2024). Hasil akhirnya disajikan dalam bentuk tabel dan visualisasi grafik yang langsung dihasilkan dari R, memberikan transparansi dan reproduktibilitas, sekaligus menjadi dasar interpretasi pola penyakit menular regional, kondisi lingkungan, kapasitas sistem kesehatan, serta strategi pengendalian yang tepat sasaran.

HASIL

Karakteristik wilayah studi disajikan berdasarkan tingkat kabupaten dan kota, meliputi 22 wilayah administratif di Provinsi Nusa Tenggara Timur. Wilayah-wilayah ini menunjukkan variasi yang cukup besar dalam jumlah kasus penyakit menular yang dilaporkan, mencerminkan kondisi epidemiologi yang tidak seragam di seluruh provinsi.

Analisis Deskriptif

Tabel 1. Statistik Deskriptif dan Uji Asumsi Untuk Kasus Penyakit Menular pada 22 Kabupaten di Provinsi Nusa Tenggara Timur

	Mean	SD	Median	Min	Max	Range
Kabupaten/Kota	11,50	6,49	11,5	1	22	21
Malaria	403,82	1175,75	24,0	2	5492	5490
TBC	280,05	184,72	271,0	41	873	832
Pneumonia	206,68	191,26	130,5	17	666	649
Kusta	20,32	21,26	14,0	0	71	71
Diare	1270,18	1318,47	931,0	0	4476	4476
DBD	176,64	189,22	136,5	3	811	808
AIDS	376,09	533,48	214,0	44	2577	2533
IMS	61,91	128,44	17,0	0	604	604

Analisis univariat mengungkapkan adanya variasi signifikan dalam kasus penyakit menular antar kabupaten dan kota di Provinsi Nusa Tenggara Timur yang menandakan beban penyakit yang berbeda-beda di tiap wilayah administratif. Diare tercatat memiliki jumlah kasus rata-rata tertinggi, sementara malaria, AIDS, dan infeksi menular seksual menunjukkan penyimpangan baku yang besar dengan nilai median lebih rendah dari rata-rata menunjukkan distribusi kasus yang terkonsentrasi di beberapa kabupaten tertentu. Tuberkulosis, pneumonia, dan demam berdarah dengue memiliki pola distribusi yang lebih stabil dengan rentang variasi yang relatif sempit antar wilayah.

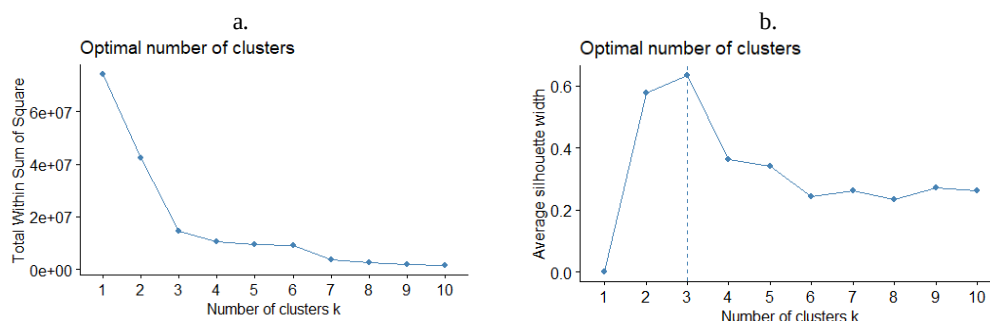
Sebelum analisis multivariat, kecukupan data dinilai terlebih dahulu untuk memastikan kesesuaiannya dengan metode klasterisasi. Nilai Kaiser Meyer Olkin secara keseluruhan mencapai 0,61 yang bermakna melebihi ambang batas minimum yang diterima sehingga menunjukkan kecukupan sampel yang memadai. Sebagian besar variabel memenuhi standar kecukupan sampel, dan uji diagnostik multikolinearitas mengonfirmasi bahwa mayoritas indikator sudah memenuhi asumsi yang diperlukan untuk analisis multivariat.

Tabel 2. Ukuran Kecukupan Sampel Kaiser-Meyer-Olkin (KMO) dan Variance Inflation Factors (VIF) Untuk Variabel Penyakit Menular Sebelum K-means Clustering

Jenis Penyakit Menular	MSA	Non Multikolinieritas
Malaria	0,31	2,482914
TBC	0,72	6,237319
Pneumonia	0,32	2,668618
Kusta	0,75	1,724782
Diare	0,69	2,804113
DBD	0,51	1,409009
AIDS	0,63	13,161558
IMS	0,64	8,917690
Overall MSA : 0,61		

Penilaian Kesesuaian Data Untuk Klasterisasi

Analisis multivariat dilakukan menggunakan algoritma klasterisasi K-Means untuk mengidentifikasi kelompok kabupaten dan kota yang memiliki profil penyakit menular serupa. Jumlah klaster optimal ditentukan dengan membandingkan metode Elbow dan Silhouette. Kedua pendekatan ini mengarah pada tiga klaster sebagai solusi terbaik, di mana metode Elbow menunjukkan penurunan jelas pada variabilitas dalam klaster hingga tiga kelompok, sementara metode Silhouette mencatat kekompakan dan pemisahan klaster maksimal pada $k=3$. Solusi klasterisasi ini dipilih untuk tahap profiling klaster selanjutnya.

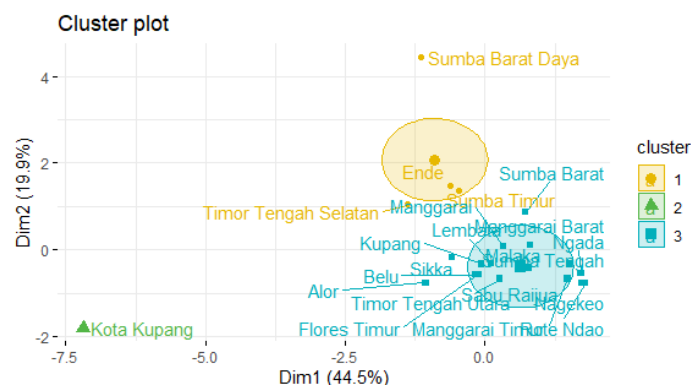


Gambar 1. Penentuan Jumlah Klaster Optimal Menggunakan a.) Metode Elbow dan b.) Metode Silhouette

Hasil Analisis Klasterisasi dan Karakteristik Klaster

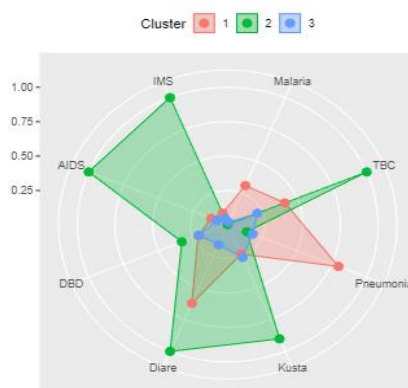
Pembagian K-means ($k=3$) menghasilkan profil berbeda yang divisualisasikan melalui *principal component* (Dim1=44,5%) dan radar plot. Klaster 1 (4 kabupaten: Sumba Timur, Timor Tengah Selatan, Ende, Sumba Barat Daya; total rata-rata=6.012) ditandai kasus malaria

tinggi (1.701), diare (2.789), dan pneumonia (535), mencerminkan hotspot hiperendemik. Klaster 2 (1 wilayah: Kota Kupang; total=8.992) didominasi AIDS (2.577), IMS (604), diare (4.476), serta tuberkulosis (873), menggambarkan sindemik perkotaan. Klaster 3 (17 kabupaten; total=1.675) memperlihatkan beban penyakit moderat merata (contoh: malaria=121, diare=724), membentuk baseline heterogen.



Gambar 2. Plot Klaster pada Komponen Utama (Dim1=44,5%) dengan District Assignment

Radar plot menonjolkan puncak malaria-diare pada Klaster 1, AIDS-IMS pada Klaster 2, serta profil seimbang pada Klaster 3. Pembagian ini selaras dengan gradien penyakit spasial yang diamati di setting serupa.



Gambar 3. Radar Plot Kasus Rata-Rata Terstandarisasi Berdasarkan Penyakit di Tiga Klaster

Tabel 3. Profil Klaster dan Rata-Rata Tingkat Kejadian Penyakit Menular di Provinsi Nusa Tenggara Timur (k=3)

Klaster	n Kabupaten/Kota	Penyakit (Mean)	Dominan	Total Mean	Interpretasi
1	4 (Sumba Timur, Timor Tengah Selatan, Ende, Sumba Barat Daya)	Malaria: 1,701; Diare: 2,789; Pneumonia: 535		6012	Hotspot malaria-diare hiperendemik
2	1 (Kota Kupang)	AIDS: 2,577; Diare: 4,476; IMS: 604; TB: 873		8992	Sindemik perkotaan (HIV-IMS-diare)
3	17 (Sumba Barat, Kupang, Timor Tengah Utara, Belu, Alor, Lembata, Flores Timur, Sikka, Ngada, Manggarai, Rote Ndao, Manggarai Barat, Sumba Tengah, Nagekeo, Manggarai Timur, Sabu Raijua, Malaka)	Malaria: 121; Diare: 724; DBD: 171		1675	Baseline heterogen moderat

PEMBAHASAN

Pembentukan tiga klaster berbeda ini menggarisbawahi beban penyakit menular di Nusa Tenggara Timur, yang dipengaruhi oleh keterkaitan faktor lingkungan, sosial-ekonomi, serta perilaku masyarakat. Pola hiperendemik di kabupaten/kota timur selaras dengan teori ekologi yang menyatakan bahwa iklim kering dan akses air terbatas memperparah penyakit yang ditularkan melalui vektor dan air, sebagaimana terbukti dari indeks parasit tahunan tinggi di provinsi Indonesia yang serupa (Guntur *et al.*, 2021). Kondisi semacam ini memberikan tantangan intervensi seragam dan lebih menguntungkan strategi lokal, serta lebih sesuai menggunakan strategi lokal yang mirip dengan strategi mitigasi wabah demam berdarah di wilayah perbatasan perkotaan-pedesaan (Kosasih *et al.*, 2016). Sindemik perkotaan di Kota Kupang mencerminkan penularan yang meningkat akibat tingkat kepadatan penduduk dan mobilitas tinggi, yang memperkuat kerangka sindemik di mana HIV/IMS terjadi bersamaan dengan penyakit diare di tengah defisit sanitasi (Fonseca-Rodríguez *et al.*, 2021). Berbeda dengan baseline pedesaan, pola ini menyerupai fenomena global di wilayah pinggiran kota di mana risiko perilaku memperburuk kerentanan, seperti terlihat pada hotspot tuberkulosis yang terkait migrasi (Aidi, 2023). Perbedaan semacam ini memperkuat model epidemiologi spasial yang menekankan determinan kontekstual dibandingkan dengan etiologi seragam (Cordes & Hennessy, 2020).

Profil moderat di sebagian besar kabupaten menunjukkan risiko yang tersebar namun dapat dimodifikasi melalui peningkatan sistemik, yang sejajar dengan klaster beban moderat dalam pemetaan kerentanan nasional (Aidi, 2023). Efektivitas K-means dalam studi ini memperluas validasi sebelumnya di kesehatan masyarakat, di mana partisi optimal berbasis silhouette mengungguli metode hierarkis pada data *noisy geospatial* (Hung *et al.*, 2022). Meskipun begitu, keterbatasan desain *cross sectional* menghalangi penetapan kausalitas; model dinamis yang menggabungkan aspek waktu (temporalitas) dapat menyempurnakan prediksi (Akinwale *et al.*, 2018). Wawasan ini memberi arahan kebijakan dengan memprioritaskan pengendalian vektor di wilayah timur, skrining sindemik di perkotaan, serta pengawasan menyeluruh di tingkat provinsi mendukung target *Sustainable Development Goal 3* di konteks kepulauan bersumber daya rendah (UNDP, 2023; Alexander *et al.*, 2021). Penelitian mendatang sebaiknya mengintegrasikan kovariat geospasial untuk memprediksi yang lebih terperinci.

KESIMPULAN

Klasterisasi K-Means ($k=3$, divalidasi silhouette) mengungkap variasi spasial kejadian penyakit menular antar kabupaten di Nusa Tenggara Timur, mencapai tujuan utama identifikasi pola di tengah pengaruh lingkungan dan sindemik. Temuan inti menunjukkan fokus hiperendemik malaria-diare di Klaster 1 wilayah timur, sindemik HIV-diare yang menonjol di Klaster 2 perkotaan (Kota Kupang), serta beban moderat yang seimbang di Klaster 3 yang luas, didukung diagnostik kuat ($KMO=0,61$). Intervensi vektor disarankan yang menyesuaikan ekologi untuk Klaster 1, pengawasan sindemik terintegrasi untuk Klaster 2, dan platform digital yang dapat diperluas di seluruh provinsi; secara teori, paradigma sindemik meningkatkan partisi multivariat; pengembangan prospektif memerlukan pemodelan spasial-temporal untuk prediksi dampak intervensi.

UCAPAN TERIMA KASIH

Peneliti menyampaikan ucapan terimakasih kepada Badan Pusat Statistik Indonesia atas dataset epidemiologi yang mudah diakses, Fakultas Kesehatan Masyarakat Universitas

Airlangga yang menyediakan sarana analisis, serta pustakawan institusi yang membantu pengelolaan referensi.

DAFTAR PUSTAKA

- Aidi, M. N. (2023) 'Province clustering based on the percentage of infectious diseases in Indonesia', *Journal of Public Health Research*, 12(3), pp. 1–10.
- Akinwale, A., Adepoju, T. & Adebisi, M. (2018) 'Clustering algorithm for a healthcare dataset using K-means', *International Journal of Computer Science and Information Technologies*, 10(2), pp. 1–5.
- Alexander, K. A., Doherty, T., Bugnon, J. *et al.* (2021) 'SDG3-related inequalities in women's, children's and adolescents' health in Latin America and the Caribbean', *BMJ Global Health*, 6(Suppl 4), e005104.
- Cordes, J. & Hennessy, J. (2020) 'Spatial analysis of COVID-19 clusters and contextual factors associated with clustering', *Spatial and Spatio-temporal Epidemiology*, 34, 100268.
- Frenk, J., Gómez-Dantés, O. & Knaul, F. M. (2017) 'The democratization of health in Mexico: financial innovations for universal coverage', *Bulletin of the World Health Organization*, 95(7), pp. 542–548.
- Guntur, R. D., Koesoemadinata, R. C., Elyazar, I. R. F. *et al.* (2021) 'Epidemiology of malaria in East Nusa Tenggara Province, Indonesia: protocol for a mixed-methods observational study', *JMIR Research Protocols*, 10(4), e23545.
- Han, J., Kamber, M. & Pei, J. (2022) *Data Mining: Concepts and Techniques*. 4th edn. San Francisco: Morgan Kaufmann.
- Hartigan, J. A. & Wong, M. A. (1979) 'Algorithm AS 136: A K-means clustering algorithm', *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1), pp. 100–108.
- Hung, P. D., Nguyen, T. H. & Tran, V. Q. (2022) 'A novel application of K-means cluster analysis in population health evaluation', *American Journal of Public Health*, 112(8), pp. 1185–1192.
- Jain, A. K. (2016) 'Data clustering: 50 years beyond K-means', *Pattern Recognition Letters*, 31(8), pp. 651–666.
- Kosasih, H., de Mast, Q., Widjaja, S. *et al.* (2016) 'The epidemiology, virology and clinical findings of dengue virus in adults living in Bandung, Indonesia', *PLoS Neglected Tropical Diseases*, 10(5), e0004390.
- MacQueen, J. (1967) 'Some methods for classification and analysis of multivariate observations', *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, pp. 281–297.
- Ministry of Health of Indonesia (2023) *Indonesia Health Profile 2023*. Jakarta: Ministry of Health of the Republic of Indonesia.
- Murray, C. J. L., Aravkin, A. Y., Zheng, P. *et al.* (2020) 'Global burden of 87 risk factors in 204 countries and territories, 1990–2019', *The Lancet*, 396(10258), pp. 1223–1249.
- Prasetyo, E., Widyastuti, R. & Nugroho, S. (2024) 'Application of K-means clustering for regional health vulnerability classification', *Pancasakti Journal of Public Health Science and Research*, 3(1), pp. 15–25.
- Putri, D. A., Sari, R. K. & Handayani, L. (2021) 'Clustering of infectious disease cases using K-means method for public health planning', *BMC Public Health*, 21, 1842.
- Rousseeuw, P. J. (1987) 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis', *Journal of Computational and Applied Mathematics*, 20, pp. 53–65.
- Statistics Indonesia (2023) *East Nusa Tenggara Province in Figures 2023*. Kupang: BPS-Statistics of East Nusa Tenggara Province.

- Suryani, D., Kusnanto, H. & Pramono, D. (2020) 'Environmental and socio-economic determinants of malaria in eastern Indonesia', *Malaria Journal*, 19, 305.
- Titaley, C. R., Ariawan, I. & Soeharno, N. (2022) 'Health service disparities and infectious disease outcomes in eastern Indonesia', *International Journal for Equity in Health*, 21, 98.
- UNDP (2023) *Goal 3: Good Health and Well-being. United Nations Development Programme.*
- World Health Organization (2022) *Global Health Estimates 2022.* Geneva: World Health Organization.