

ALGORITMA K-MEANS UNTUK PENGELOMPOKAN LOKASI SUMUR BERDASARKAN TINGKAT KONTAMINASI BAKTERI COLIFORM DI KOTA CIMAH

Ade Rizki AS^{1*}, Novita Lestari Anggreini²

¹*Program Studi Teknik Informatika, ²Politeknik TEDC Bandung*

*Penulis Korespondensi: aderizkias0@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan lokasi sumur di Kota Cimahi berdasarkan tingkat kontaminasi bakteri *coliform* dengan menggunakan algoritma *K-Means Clustering*. Kualitas air merupakan isu krusial dalam pembangunan kota terutama di daerah padat penduduk seperti Cimahi. Melalui teknik pengelompokan data (*clustering*), penelitian ini mengklasifikasikan data sumur ke dalam empat kategori yaitu tingkat kontaminasi rendah, sedang, tinggi dan sangat tinggi. *Dataset* yang digunakan bersumber dari Dinas Lingkungan Hidup Kota Cimahi tahun 2018–2022 yang tersedia di *open data* Cimahi. Data dianalisis menggunakan *Google Colab* dan dilakukan normalisasi sebelum dilakukan proses *clustering*. Hasil pengelompokan menunjukkan pemisahan cluster yang cukup baik dengan *nilai silhouette score* sebesar 0,82. Temuan ini diharapkan dapat membantu pemerintah daerah dalam mengidentifikasi wilayah yang berisiko tinggi terhadap pencemaran air tanah dan mengambil langkah-langkah mitigasi yang tepat.

Keywords: Cimahi, *Coliform*, *K-Means*, Kualitas Air, *Clustering*

PENDAHULUAN

Kota Cimahi yang berada di Provinsi Jawa Barat merupakan salah satu wilayah dengan luas terkecil di Indonesia namun memiliki tingkat kepadatan penduduk yang cukup tinggi. Pada tahun 2001 Cimahi resmi menjadi kota otonom. Sejak saat itu Cimahi menunjukkan perkembangan yang pesat di berbagai sektor seperti infrastruktur, pendidikan dan industri.

Namun kemajuan tersebut juga membawa dampak lingkungan yang cukup serius termasuk pada kualitas air sumur sebagai sumber utama air bersih bagi masyarakat. Pertumbuhan populasi dan industrialisasi meningkatkan kebutuhan lahan yang berdampak pada berkurangnya kemampuan tanah menyerap air hujan. Selain itu aktivitas domestik dan limbah industri turut berkontribusi terhadap meningkatnya tingkat pencemaran air sumur.

Kualitas air harus memenuhi standar tertentu untuk memastikan konsumsi yang aman. Air bersih adalah air yang memenuhi standar tertentu dan dapat digunakan untuk kebutuhan sehari-hari (Djana 2023). Kandungan zat besi (Fe^{3+}) yang berlebihan pada air jika dikonsumsi dalam jumlah banyak dapat menyebabkan gangguan *metabolisme* (Kristiandi, Anisah, and Junardi 2023). Selain

itu keberadaan bakteri *coliform* menjadi indikator utama kebersihan air. Bakteri *coliform* merupakan kelompok bakteri yang dapat digunakan sebagai indikator pencemaran air limbah dan buruknya kebersihan air. terdapat *bakteri coliform* yang bersifat *patogen* yaitu bakteri yang dapat menimbulkan penyakit (Ayer et al. 2023). Bakteri ini terbagi menjadi dua kelompok *fecal* dan *non fecal*(total), *coliform fekal* yang berasal dari tinja manusia dan hewan termasuk *E. coli* serta *coliform non fekal* yang berasal dari hewan atau tumbuhan mati (Agustina et al. 2024). Air minum yang memenuhi standar kesehatan harus bebas bakteri tersebut untuk mencegah risiko penyakit pencernaan.

Pemetaan tingkat kontaminasi air sumur di Kota Cimahi menjadi langkah yang sangat penting untuk mengidentifikasi masalah dan merancang solusi. Menurut data dari Dinas Lingkungan Hidup tahun 2019 kualitas air tanah di sejumlah lokasi memerlukan pengawasan yang lebih ketat serta upaya pengelolaan kualitas air yang lebih efektif (Putri and Wardhani 2021).

Salah satu metode yang dapat diterapkan untuk mengelompokkan data kualitas air adalah algoritma *K-Means*

Clustering. Algoritma adalah ilmu yang menangani penyelesaian suatu masalah melalui serangkaian langkah spesifik yang kemudian disusun secara sistematis dan dalam bahasa yang logis serta memiliki tujuan (Maulana Hasan and Yahfizham 2023). Algoritma ini berfungsi dengan mengelompokkan data berdasarkan kesamaan tertentu seperti tingkat kontaminasi. Tujuan utamanya adalah membagi data atau objek ke dalam beberapa *cluster* di mana setiap *cluster* berisi data yang memiliki kemiripan satu sama lain (Mirantika 2021). Algoritma *K-Means* dapat digunakan untuk memetakan lokasi-lokasi dengan tingkat pencemaran yang serupa mempermudah pemerintah dalam mengambil langkah intervensi yang tepat dan efisien. Pendekatan ini memungkinkan analisis data yang terstruktur dan memberikan hasil yang jelas serta mudah diinterpretasikan untuk mendukung pengambilan keputusan yang lebih baik.

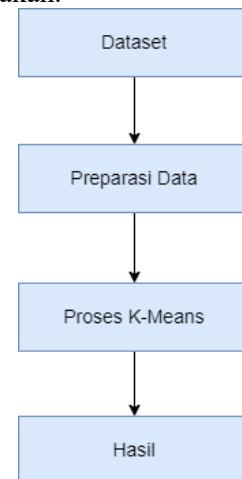
Dengan memanfaatkan algoritma *K-Means Clustering* tingkat kontaminasi air sumur di Kota Cimahi dapat dipetakan berdasarkan konsentrasi *coliform*. *Coliform* yang merupakan indikator utama pencemaran bakteri dalam air berperan penting dalam menentukan kelayakan air sumur untuk digunakan sehari-hari ataupun dikonsumsi.

Algoritma *K-Means* dapat digunakan untuk pengelompokan wilayah berdasarkan tingkat kontaminasi *coliform* ke dalam kategori rendah, sedang, tinggi dan sangat tinggi. Hasil dari pengelompokan ini membantu mengidentifikasi wilayah-wilayah kritis secara lebih efektif sehingga mendukung proses pengambilan keputusan untuk pengelolaan kualitas air sumur. Dengan demikian penerapan algoritma *K-Means* untuk memetakan kontaminasi *Coliform* memberikan solusi strategis dalam menghadapi pemetaan air bersih di Kota Cimahi.

METODOLOGI PENELITIAN

Metodologi penelitian merupakan aspek penting dalam desain dan implementasi penelitian (Sina, 2024). Penelitian empiris adalah yang datanya tersedia dalam bentuk angka (Syahrizal and Jailani 2023). Penelitian ini menggunakan pendekatan kuantitatif dengan metode data *mining* menggunakan

algoritma *K-means clustering*. Tujuan utamanya adalah mengelompokkan lokasi sumur di Kota Cimahi berdasarkan tingkat kontaminasi *coliform*. Metode yang diusulkan merupakan suatu proses untuk memudahkan penelitian dapat dilihat pada gambar 1 Metode yang diusulkan.



Gambar 1. Metode yang diusulkan

Dataset

Dataset adalah kumpulan data yang dapat digunakan sebagai bahan percobaan penelitian (Yuliska and Syaliman 2020). Dalam penelitian ini penulis menggunakan data lokasi sumur yang terkontaminasi oleh bakteri *coliform* di Kota Cimahi. Data yang dikumpulkan merupakan data dari tahun 2018-2022 yang berjumlah 326 data. Data sebelum di preparasi yang memiliki 13 kolom (id, kode provinsi, nama, provinsi, kode kabupaten/kota, nama kabupaten/kota, lokasi sumur, waktu sampling, titik koordinat lintang, titik koordinat bujur, *coliform*, jumlah, satuan, tahun)

Preparasi Data

Preparasi data dimaksudkan untuk memastikan bahwa data yang akan digunakan memiliki kualitas optimal, bebas kesalahan, terorganisir dengan baik dan siap untuk analisis lebih lanjut (Buntara, Napitupulu, and Gusriani 2023). Setelah ditelaah kembali data lokasi sumur ternyata 2 baris data adalah 1 data karena 1 baris ada kolom *fecal coliform* dan 1 baris kolom total *coliform* maka dari itu 2 baris itu disatukan dalam satu baris dengan membuat kolom baru yaitu *fecal coliform* dan total *coliform*. Preparasi yang dilakukan

penulis terdapat 4 kolom data yang akan digunakan (lokasi sumur, tahun, *fecal coliform*, total *coliform*). Berikut adalah tabel 1 Data yang telah di preparasi.

Tabel 1 Data yang telah dipreparasi			
Lokasi Sumur	Tahun	Fecal Coliform	Total Coliform
Air sumur rumah hj.			
Kustiah rt.08/01 kel. Cimahi kec. Cimahi tengah	2018	14	109
...
02/16 Cibabat	2022	82970	87040

Proses K-Means

Pengelompokan *K-means* merupakan salah satu teknik pengelompokan yang paling umum digunakan dalam analisis data. Algoritma ini bekerja dengan membagi data ke dalam K kelompok yang telah ditentukan sebelumnya di mana K adalah jumlah kelompok yang diinginkan. Proses ini dimulai dengan memilih secara acak K pusat *cluster* (*centroid*) di ruang data dan kemudian mengelompokkan setiap titik data ke dalam kelompok dengan *centroid* terdekat (*inertia*) (Yudhistira and Andika 2023). Data sebelumnya yang telah di preparasi selanjutnya akan dilakukan proses di *Google Colab* menggunakan algoritma *K-means Clustering*. Berikut langkah-langkah yang dilakukan:

Pemilihan Variabel

Variabel yang digunakan untuk pengelompokan adalah kolom *fecal coliform* dan total *coliform*, yang menunjukkan jumlah bakteri *coliform* dalam satuan jumlah/1000 ml.

Penentuan Jumlah Cluster

Jumlah *cluster* yang digunakan adalah empat kategori yaitu rendah, sedang, tinggi dan sangat tinggi tingkat kontaminasi *coliform*.

Normalisasi Data dengan Standardscaler

Normalisasi diperlukan karena nilai untuk *fecal coliform* dan total *coliform* sangat berbeda. Jika normalisasi tidak dilakukan hasil pengelompokan mungkin condong ke nilai yang lebih tinggi.

Standardscaler bekerja dengan rumus:

$$X' = \frac{x - \mu}{\sigma}$$

dimana:

X' = nilai setelah dinormalisasi

X = nilai asli,

μ = rata-rata seluruh data di kolom tersebut,

σ = standar deviasi seluruh data dikolom tersebut

Implementasi K-Means

Algoritma *K-Means* diterapkan menggunakan pustaka *scikit-learn* di *Google Colab*. Parameter yang digunakan adalah *n_clusters*=4 untuk menghasilkan empat kelompok.

Hasil Clustering

Berdasarkan hasil *clustering* lokasi sumur dikategorikan 0, 1, 2 dan 3. Dimana cluster 0 adalah sumur dengan tingkat kontaminasi rendah lalu cluster 1 yang dikategorikan tingkat kontaminasi sedang, cluster 2 yang dikategorikan tingkat kontaminasi tinggi dan 3 kategori sangat tinggi.

Analisis

Analisis dilakukan untuk mengevaluasi kinerja algoritma *K-Means* dalam mengklasifikasikan lokasi sumur berdasarkan tingkat kontaminasi *coliform* di Kota Cimahi. Evaluasi dilakukan dengan mengamati sebaran cluster dan kestabilan hasil cluster. Metrik evaluasi yang digunakan meliputi nilai *inertia* dan *silhouette Score* untuk mengukur kualitas *clustering*.

HASIL DAN PEMBAHASAN

Hasil penelitian membahas proses pengolahan data, proses pengujian serta implementasi algoritma *K-Means Clustering*.

Data Yang Digunakan

Data yang digunakan untuk dianalisis adalah data dari Dinas Lingkungan Hidup Kota Cimahi dari tahun 2018-2022. Sebelum digunakan data telah melalui proses pembersihan untuk menghapus atribut yang tidak diperlukan dan melakukan penyesuaian kolom.

Selanjutnya dilakukan dengan menambahkan kolom *cluster* dimana untuk memberikan label lokasi sumur berdasarkan tingkat kontaminasi *coliform*. Tabel 2 adalah *Dataset* dengan Kolom *Cluster*.

Tabel 2 *Dataset* dengan Kolom *Cluster*

Lokasi Sumur	Tahun	Fecal Coliform	Total Coliform	cluster
Air sumur rumah hj. Kustiah rt.08/01 kel.	2018	14	109	
Cimahi kec. Cimahi tengah				
...	
02/16 Cibabat	2022	82970	87040	

Pengolahan Data

Pada tahap ini dilakukan pengolahan data dengan tujuan untuk mempersiapkan *dataset* agar dapat digunakan dengan lebih baik dalam proses *clustering* menggunakan algoritma *K-Means*. Langkah pertama yang dilakukan adalah mengisi nilai yang kosong dengan angka 0 dan membuat kolom *fecal* dan *total coliform* menjadi satu baris dapat dilihat pada gambar 2 Proses Pengolahan Data.

```
① df[“jumlah”] = df[“jumlah”].fillna(0)

df_pivot = df.pivot_table(index=[“lokasi_sumur”, “tahun”], columns=“coliform”)
df_pivot.head()

② coliform lokasi_sumur tahun FECAL COLIFORM TOTAL COLIFORM
   0 01/01 Baros 2022 2160.0 37849.0
   1 01/02 Cibeber 2022 100.0 2110.0
   2 01/08 Cimahi 2022 6370.0 21420.0
   3 01/09 Karangmekar 2022 310.0 8450.0
   4 01/12 Citeureup 2022 310.0 10760.0
```

Gambar 2. Proses Pengolahan Data

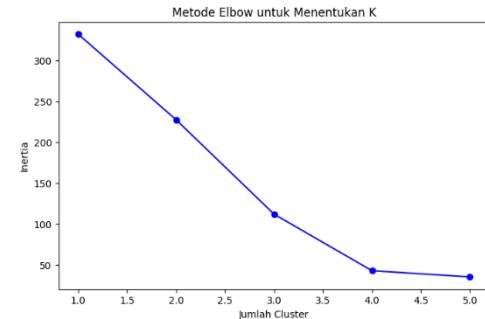
Lalu data dinormalisasi menggunakan metode *StandardScaler*. Normalisasi diperlukan karena data yang digunakan memiliki skala nilai yang berbeda sehingga perlu diubah agar semua fitur memiliki distribusi yang sebanding. Gambar 3 adalah proses Normalisasi Data Dengan Metode *Standardscaler*.

```
scaler = StandardScaler()
df_pivot[“FECAL COLIFORM”, “TOTAL COLIFORM”] =
scaler.fit_transform(df_pivot[“FECAL COLIFORM”, “TOTAL COLIFORM”])
```

Gambar 3. Normalisasi Data Dengan Metode *Standardscaler*.

Setelah data di normalisasi langkah berikutnya yang dilakukan adalah menentukan jumlah *cluster* yang optimal menggunakan *elbow method* yang mengevaluasi nilai *inertia* (jumlah jarak

antara titik data dan pusat *cluster*). Gambar 4 adalah *Elbow Method*.



Gambar 4. *Elbow Method*

Dari grafik yang dihasilkan nilai *K* optimal ditentukan berdasarkan titik siku pada diagram. Dalam penelitian ini *K* = 4 dipilih berdasarkan kategori pencemaran yaitu 0 adalah Pencemaran rendah, 1 adalah Pencemaran sedang, 2 adalah Pencemaran tinggi dan 3 adalah Pencemaran sangat tinggi.

Implementasi Algoritma *K-Means*

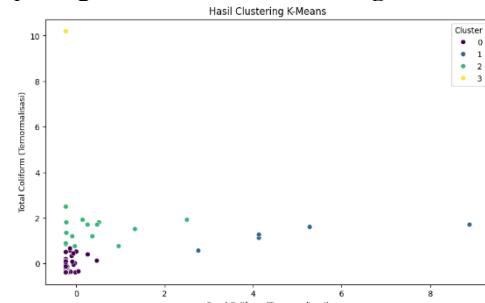
Setelah pengolahan data dilakukan dari mengisi nilai yang kosong sampai menentukan *cluster* selanjutnya akan dilakukan implementasi algoritma *K-Means* untuk mengelompokan lokasi sumur berdasarkan tingkat kontaminasi *coliform*. Gambar 5 adalah *Code K-Means*.

```
① kmeans = KMeans(n_clusters=4, random_state=42)
df_pivot[“Cluster”] = kmeans.fit_predict(df_pivot[“FECAL COLIFORM”, “TOTAL COLIFORM”])
df_pivot.head()

② coliform lokasi_sumur tahun FECAL COLIFORM TOTAL COLIFORM Cluster
   0 01/01 Baros 2022 0.004627 0.520717 0
   1 01/02 Cibeber 2022 -0.221874 -0.339000 0
   2 01/08 Cimahi 2022 0.467526 0.125510 0
   3 01/09 Karangmekar 2022 -0.198784 -0.186489 0
   4 01/12 Citeureup 2022 -0.198784 -0.130921 0
```

Gambar 5. *Code K-Means*

Hasil dari proses ini adalah bahwa setiap sumur diklasifikasikan menjadi salah satu dari empat cluster tergantung pada tingkat kontaminasi *coliform* dan dapat dilihat pada gambar 6 Hasil *Clustering K-Means*.



Gambar 6. Hasil *Clustering K-Means*

Hasil Clustering

Setelah melalui proses *K-Means clustering* didapat data bahwa ada 136 data untuk cluster 0 yang artinya rendah, 5 data untuk cluster 1 yang artinya sedang, 15 data untuk cluster 2 yang artinya tinggi dan 1 data untuk cluster 3 yang artinya sangat tinggi. Dapat dilihat pada gambar 7 Jumlah Data Per Cluster.

Cluster	Jumlah Data
0	145
1	15
2	5
3	1

Gambar 7. Jumlah Data per Cluster

Lalu untuk mengetahui keakuratan *K-means* untuk *clustering* dapat dilihat pada gambar 8 Nilai Akurasi Silhouette Score yaitu mengukur seberapa baik data dikelompokkan dengan mempertimbangkan jarak antar *cluster*, *Silhouette Score* berkisar antara -1 sampai 1.

```
▶ from sklearn.metrics import silhouette_score
  sil_score = silhouette_score(df_pivot)
  print("Silhouette Score:", sil_score)
```

→ Silhouette Score: 0.8063046264100807

Gambar 8. Nilai Akurasi Silhouette Score

Sementara itu nilai *inertia* yang jika nilainya semakin rendah maka menunjukkan bahwa jarak antara data individual dan pusat *cluster* (*centroid*) juga kecil. Gambar 9 adalah Nilai *Inertia* dari *K-Means* untuk pengelompokan lokasi sumur berdasarkan kontaminasi *coliform*.

```
▶ # Menampilkan nilai inertia
  print("Nilai Inertia:", kmeans.inertia_)
```

→ Nilai Inertia: 43.07981539438079

Gambar 9. Nilai Inertia

Analisis Karakteristik Setiap Cluster

Dari hasil pengelompokan menggunakan algoritma *K-Means*, lokasi sumur dikelompokkan menjadi tiga *cluster* berdasarkan tingkat pencemaran *fecal coliform* dan total *coliform*. Karakteristik masing-masing klaster adalah sebagai berikut:

Cluster 0 (kontaminasi rendah), mayoritas sumur dalam *cluster* ini memiliki nilai *fecal coliform* dan total *coliform* yang rendah atau mendekati nol setelah

dinormalisasi. Air sumur dalam kategori ini lebih aman untuk dikonsumsi dibandingkan *cluster* lainnya.

Cluster 1 (Kontaminasi Sedang), sumur dalam cluster ini memiliki tingkat *coliform* yang lebih tinggi dibandingkan *cluster 0* tetapi masih dalam batas wajar untuk air tanah. Air dari sumur ini memerlukan pengolahan lebih lanjut sebelum digunakan untuk konsumsi.

Cluster 2 (Kontaminasi Tinggi), *cluster* ini terdiri dari sumur dengan tingkat *coliform* yang tinggi. Air sumur dari cluster ini tidak layak dikonsumsi tanpa pengolahan khusus karena dapat menyebabkan berbagai penyakit. Tabel 3 adalah contoh data yang telah dianalisis dari berbagai *cluster*.

Cluster 3 (Kontaminasi Sangat Tinggi), *cluster* ini terdiri dari sumur dengan tingkat *coliform* yang sangat tinggi jauh dari kelayakan konsumsi maupun pemakaian sehari hari. Air dari sumur-sumur ini memerlukan tindakan pengolahan yang intensif serta pemantauan berkala. Tabel 3 adalah Data Yang Telah Dianalisis.

Tabel 3 Data Yang Telah Dianalisis.

Lokasi Sumur	Tahun	Fecal Coliform	Total Coliform	cluster
01/01 Baros	202 2	0.0046273985 95781799	0.5207171195 610736	0
03/02 Setiamanah	202 2	4.1300345186 43830	1.1242923309 26240	1
03/07 Cibeber	202 2	25.038.412.53 5.715.800	19.210.087.23 2.746.700	2
Jl. Kolonel Masturi Gang Margaresek 2 RT 02 RW 10 Kel. Cimahi	202 0	- 0.2328693865 6689926	10.194.639.03 6.810.900	3

Evaluasi Model Clustering

Berdasarkan hasil evaluasi model, diperoleh nilai *silhouette score* sebesar 0,806 dan nilai *inertia* sebesar 43,08. Nilai *silhouette score* yang mendekati 1 menunjukkan kualitas hasil pengelompokan data dengan menggunakan algoritma *K-Means* sangat baik. Data individu relatif dekat dengan *cluster* nya sendiri dan jauh dari *cluster* lainnya. Hal ini menunjukkan struktur cluster yang kompak dan terpisah dengan jelas.

Nilai *inertia* yang relatif rendah menunjukkan jarak antara data individu dengan pusat *cluster* nya (*centroid*) juga kecil. Hal ini menunjukkan cluster yang terbentuk secara optimal dan efisien mengelompokkan data berdasarkan kedekatan.

Dengan demikian, dapat disimpulkan bahwa algoritma *K-Means* cukup efektif dalam mengelompokkan lokasi sumur di Kota Cimahi berdasarkan tingkat pencemaran *coliform*.

KESIMPULAN

Algoritma *K-Means* terbukti mampu mengelompokkan lokasi sumur berdasarkan tingkat kontaminasi *coliform*. Pengelompokan yang dilakukan berhasil mengelompokkan sumur ke dalam beberapa kategori berdasarkan tingkat kontaminasi yaitu rendah, sedang dan tinggi. Hal ini memudahkan identifikasi lokasi dengan risiko kontaminasi yang lebih tinggi. Nilai *inertia* sebesar 43,08 menunjukkan bahwa pemisahan cluster cukup baik meskipun masih dapat ditingkatkan dengan menyelidiki angka K yang lebih optimal. Selanjutnya *Silhouette score* menunjukkan bahwa sebagian besar data dikelompokkan dengan benar meskipun ada beberapa data yang terletak di perbatasan antar *cluster*. Hasil pengelompokan ini dapat digunakan sebagai acuan bagi pihak-pihak terkait untuk melakukan tindakan pencegahan dan meningkatkan kualitas air di daerah dengan tingkat kontaminasi yang tinggi. Untuk penelitian lebih lanjut beberapa saran dapat dipertimbangkan untuk meningkatkan akurasi dan mendapatkan hasil yang lebih baik. Selain parameter *coliform* faktor lingkungan lain seperti curah hujan, jenis tanah, polusi industri, dan sistem drainase di sekitar sumur juga dapat dimasukkan dalam penelitian. Hal ini dapat menghasilkan wawasan yang lebih komprehensif dalam memahami pola polusi. Untuk memperoleh hasil yang lebih bermakna peneliti disarankan untuk menggunakan visualisasi berbasis peta sehingga distribusi *cluster* dapat divisualisasikan secara geografis dan pembaca serta pembuat kebijakan dapat lebih mudah memahami area yang rawan kontaminasi.

DAFTAR PUSTAKA

Agustina, Natasha, Ika Afifah Nugraheni, Ainun Naim, Fakultas Sains, and Universitas Aisyiyah Yogyakarta. 2024. “Analisis Kualitas Mikrobiologis Air Sungai Melalui Deteksi Total Coliform Dan Escherichia Coli Menggunakan Metode Most Probable Number (MPN) Abstrak Analysis of Microbiological Quality of River Water through Detection of Total Coliform and Escherichia C.” 2(September):1521–34.

Ayer, Popi Ida Laila, Vera K. Mandey, Korinus Rejauw, Ervina Indrayani, Albida Rante Tasak, and Tien Nova B. Yenusi. 2023. “Kepadatan Bakteri Coliform Sebagai Indikator Pencemaran Biologis Di Perairan Wisata Enggros, Kota Jayapura, Papua.” *Jurnal Ilmu Kelautan Kepulauan* 6(2):904–16. doi: 10.33387/jikk.v6i2.7366.

Buntara, Michelle Selina, Herlina Napitupulu, and Nurul Gusriani. 2023. “Pemrograman Python Untuk Peramalan Data Deret Waktu Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (Sarima).” *In Search* 22(2):354–62. doi: 10.37278/insearch.v22i2.774.

Djana, Miftahul. 2023. “Analisis Kualitas Air Dalam Pemenuhan Kebutuhan Air Bersih Di Kecamatan Natar Hajimena Lampung Selatan.” *Jurnal Redoks* 8(1):81–87. doi: 10.31851/redoks.v8i1.11853.

Kristiandi, Kiki, Susi Anisah, and Junardi. 2023. “Analisis PH, Kadar Besi Dan Kandungan Total Coliform Pada Air Minum Isi Ulang Galon.” *Seminar Nasional Terapan Riset Inovatif* 9(2):139–46.

Maulana Hasan, and Yahfizham. 2023. “Pengenalan Algoritma Pada Pembelajaran Pemrograman Komputer.” *Comit: Communication, Information and Technology Journal* 2(2):285–99. doi:

- 10.47467/comit.v2i2.1386.
- Mirantika, Nita. 2021. "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 Di Provinsi Jawa Barat." *Nuansa Informatika* 15(2):92–98. doi: 10.25134/nuansa.v15i2.4321.
- Putri, L. O. L., and E. Wardhani. 2021. "Analysis of Groundwater Quality in Cimahi City of West Java Province." *IOP Conference Series: Earth and Environmental Science* 894(1). doi: 10.1088/1755-1315/894/1/012037.
- Sina, I. 2024. "Metodologi Penelitian Kualitatif Dan Kuantitatif Untuk Ilmu Sains." 63.
- Syahrizal, Hasan, and M. Syahran Jailani. 2023. "Jenis-Jenis Penelitian Dalam Penelitian Kuantitatif Dan Kualitatif." *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora* 1(1):13–23. doi: 10.61104/jq.v1i1.49.
- Yudhistira, Aditia, and Rio Andika. 2023. "Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering." *Journal of Artificial Intelligence and Technology Information (JAITI)* 1(1):20–28. doi: 10.58602/jaiti.v1i1.22.
- Yuliska, Yuliska, and Khairul Umam Syaliman. 2020. "Literatur Review Terhadap Metode, Aplikasi Dan Dataset Peringkasan Dokumen Teks Otomatis Untuk Teks Berbahasa Indonesia." *IT Journal Research and Development* 5(1):19–31. doi: 10.25299/itjrd.2020.vol5(1).4688.