

CLUSTERING DAERAH TERDAMPAK SAMPAH DI INDONESIA MENGUNAKAN ALGORITMA DBSCAN.

Dessy Santi¹, Wulan Maharani², Syahrullah³, Deny Wiria Nugraha⁴, Baso Mukhlis⁵, Agustinus Kali⁶

Teknik Informatika, Universitas Tadulako^{1,2,3,4}

Teknik Elektro, Universitas Tadulako^{5,6}

dessy@untad.ac.id¹, wulanmhrn@gmail.com², syahrullah@untad.ac.id³, deny.wiria.nugraha@untad.ac.id⁴

ABSTRACT

The waste problem in Indonesia is a complex and evolving environmental issue, particularly in areas with high population density and economic activity. This study aims to cluster regions affected by waste issues using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. DBSCAN was chosen for its ability to identify spatial patterns and detect outliers without requiring a predefined number of clusters. The data used includes spatial and non-spatial information related to waste volume and regional characteristics across various provinces in Indonesia. The results show that DBSCAN effectively groups waste-affected areas into several clusters based on data density and spatial proximity. These clusters can serve as a foundation for determining policy priorities for regional and national waste management. This research is expected to contribute to the development of more targeted and data-driven waste management strategies.
Keywords : *Clustering, DBSCAN, Outlier, Wasted, Spatial*

INTISARI

Permasalahan sampah di Indonesia merupakan isu lingkungan yang kompleks dan terus berkembang, terutama di daerah-daerah dengan tingkat kepadatan penduduk dan aktivitas ekonomi yang tinggi. Penelitian ini bertujuan untuk mengelompokkan daerah-daerah yang terdampak oleh permasalahan sampah menggunakan algoritma *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN). Algoritma ini dipilih karena kemampuannya dalam mengidentifikasi pola spasial serta mendeteksi *outlier* tanpa memerlukan jumlah kluster yang telah ditentukan sebelumnya. Data yang digunakan mencakup informasi spasial dan non-spasial terkait volume sampah dan karakteristik wilayah di berbagai provinsi di Indonesia. Hasil penelitian menunjukkan bahwa DBSCAN mampu mengelompokkan daerah-daerah terdampak sampah ke dalam beberapa *cluster* berdasarkan kepadatan data dan kedekatan spasialnya. *Cluster* tersebut dapat digunakan sebagai dasar dalam penentuan prioritas kebijakan pengelolaan sampah di tingkat regional maupun nasional. Penelitian ini diharapkan dapat memberikan kontribusi dalam penyusunan strategi penanganan sampah yang lebih tepat sasaran dan berbasis data.

Kata kunci: *Clustering, DBSCAN, Outlier, Sampah, Spasial.*

I. PENDAHULUAN

Menurut *World Health Organization* (WHO), sampah adalah sesuatu yang tidak digunakan, tidak terpakai, tidak disukai atau dibuang yang timbul akibat dari aktivitas manusia dan tidak terjadi dengan sendirinya (Asriati & Harjanti, 2023). Tingginya jumlah penduduk mempengaruhi tingkat konsumsi dari masyarakat yang menyebabkan jumlah volume sampah yang dihasilkan pun bertambah dan karena tingginya volume sampah tersebut dapat mengakibatkan berbagai kerugian dalam berbagai aspek seperti Kesehatan dan juga lingkungan .

Sebagai salah satu negara berkembang dengan jumlah penduduk lebih dari 200 juta jiwa, masalah kesehatan lingkungan di Indonesia menjadi sangat kompleks terutama di kota-kota besar.

Indonesia adalah negara penghasil sampah terbesar ke-5 di dunia pada 2020 dalam laporan Bank Dunia yang bertajuk *The Atlas of Sustainable Development Goals 2023*. Menurut laporan tersebut, pada 2020 Indonesia memproduksi sekitar 65, 2 juta ton sampah. Negara-negara dengan produksi sampah tinggi biasanya memiliki populasi penduduk yang besar. Namun, volume

sampah suatu negara tak semata-mata ditentukan dari jumlah warga, melainkan juga dari gaya hidup masyarakatnya. Produksi sampah juga erat kaitannya dengan pendapatan negara. Negara-negara yang lebih kaya memiliki tingkat konsumsi lebih tinggi, dan karenanya menghasilkan lebih banyak limbah. Menurut data Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) Kementerian Lingkungan Hidup dan Kehutanan (KLHK), Indonesia menghasilkan 19,45 juta ton timbunan sampah pada 2022 dengan jumlah kepadatan penduduk Indonesia mencapai 275,77 juta orang Menurut laporan Badan Pusat Statistik (BPS). Berdasarkan provinsi, timbunan sampah terbanyak berasal dari Jawa Tengah, volumenya mencapai 4,25 juta ton atau 21,85% dari total timbunan sampah nasional tahun lalu. DKI Jakarta berada di peringkat kedua dengan volume timbunan sampah 3,11 juta ton (15,98%). Lalu Jawa Timur menempati peringkat ketiga dengan volume timbunan sampah 1,63 juta ton (8,38%), diikuti Jawa Barat 1,11 juta ton (5,7%) dan Riau 1,05 juta ton (5,39%). Adapun provinsi penghasil sampah paling sedikit pada 2022 adalah Papua Barat, yakni hanya 18,16 ribu ton. Di atas Papua Barat, ada Kalimantan Utara dan Nusa Tenggara Barat yang menghasilkan timbunan sampah masing-masing 20,75 ribu ton dan 37,44 ribu ton.. Pada tahun 2022, timbunan sampah di Indonesia mencapai 36,218,012.28 ton/tahun yang mana pengurangan sampah terjadi sebanyak 14.88% atau 5,390,999.15 ton/tahun dan penanganan sampah sebanyak 49.12% atau 17,791,815.35 ton/tahun. Sedangkan sampah terkelola sebanyak 64.01% atau 23,182,814.50 ton/tahun dan sampah tidak terkelola sebanyak 35.99% atau 13,035,197.78 ton/tahun.

Jika diamati lebih jauh terdapat permasalahan mengenai pengolahan sampah di Indonesia, Yang dimana dalam pengolahan sampah pemerintah memiliki kapasitas yang

rendah dan juga kurangnya kepedulian terhadap masyarakat Indonesia terhadap lingkungan serta kurangnya rasa tanggung jawab dari industri-industri yang ada. Pengelolaan sampah yang terjadi selama ini belum sesuai dengan metode dan teknik pengelolaan sampah yang ramah lingkungan sesuai dengan ketentuan yang ada. Sampah yang tidak dikelola dengan baik telah menjadi masalah nasional sehingga pengelolaannya harus dikelola secara komprehensif dan terintegrasi dari hulu hingga hilir agar seringkali bermanfaat secara ekonomi, menyehatkan masyarakat, aman bagi lingkungan, dan dapat mengubah perilaku Masyarakat. Oleh karena itu, untuk dapat menanggulangi permasalahan sampah yang ada. Serta dibutuhkan sebuah solusi yang tepat dan dapat mendukung terselesainya permasalahan tersebut. Yang dimana dalam persebaran sampah terdapat ketidakpastian mengenai pola persebaran sampah di berbagai wilayah di Indonesia. pengelompokkan daerah dengan karakteristik serupa dapat membantu pemerintah, Lembaga lingkungan, dan masyarakat untuk merancang strategi pengolahan sampah yang lebih tepat dan efektif. Klasterisasi dapat membantu memfasilitasi pemahaman yang mendalam mengenai perbedaan regional dan kebutuhan spesifik pada masing- masing wilayah.

II. LANDASAN TEORI

A. *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*

DBSCAN merupakan algoritma *clustering* yang dapat menemukan area berdasarkan kerapatan yang tinggi untuk membentuk *cluster*. *Cluster* pada algoritma DBSCAN didefinisikan sebagai area yang mempunyai sampel yang padat atau kerapatan yang tinggi dan dibedakan dari area kerapatan yang rendah (*noise*) (Indini et al., 2022). Dalam algoritma DBSCAN terdapat 2

parameter yang digunakan, ialah *Eps* (*Epsilon*) dan *MinPts* (*Minimum Points*). Yang Dimana *Eps* merupakan jarak maksimum antar titik agar dianggap sebagai tetangga. Dan *MinPts* merupakan jumlah minimum titik dalam *radius* agar dianggap sebagai *Core Point*. Adapun Langkah-langkah kerja dari algoritma DBSCAN ialah :

1. Menentukan nilai parameter *Eps* dan *MinPts*
2. Menentukan nilai *p* secara *random* sebagai titik awal
3. Menghitung *Eps* (*Radius/Jarak*) dari semua titik terhadap *p*
4. *Cluster* akan terbentuk ketika titik titik-titik memiliki *radius Eps* dan bernilai lebih dari *MinPts* dan titik *P* sebagai *Core Point*.
5. Ulangi Langkah 3 dan 4 pada seluruh titik. Jika titik *P* merupakan titik *Border* dan tidak ada titik yang *density reachable* terhadap *p*, maka dilakukan dilanjutkan terhadap titik lain.

B. Clustering Dalam Data Mining

Clustering adalah suatu teknik dalam data *mining* yang bertujuan untuk mengelompokkan data- data (objek) ke dalam beberapa *cluster* atau kelompok sehingga objek yang serupa disatukan kedalam *cluster* yang sama, dan yang berbeda harus menjadi bagian dari *cluster* yang berbeda (Adha et al., 2021). Dalam konteks pengolahan sampah, *clustering* dapat membantu dalam pengelompokkan daerah-daerah berdasarkan Tingkat kepadatan sampahnya sehingga diperoleh daerah-daerah mana saja yang terdampak sampah di Indonesia dan hal tersebut dapat memudahkan proses analisis.

C. Epsilon dan Minimum Points (*MinPts*)

Epsilon merupakan batas maksimum antar pengguna dan *MinPts* batas minimum jumlah pengguna dalam sebuah *cluster*. Nilai *epsilon* yang semakin besar tidak selalu membuat nilai *modularity* semakin tinggi, dikarenakan pengaruh nilai *similarity* dengan nilai *epsilon*.

D. Silhouette Score

Silhouette Score merupakan suatu metrik yang digunakan untuk mengukur kecocokan objek dalam suatu *cluster* dibandingkan dengan *cluster* lainnya. Yang dimana rentang nilai yang dihasilkan adalah antara -1 sampai dengan 1, skor yang mendekati 1 menunjukkan bahwa titik data tersebut terkelompok dengan baik, dan skor yang mendekati -1 menunjukkan bahwa titik data tersebut salah diklasifikasikan dan seharusnya termasuk dalam *cluster* tetangga. Sedangkan skor 0 menunjukkan bahwa titik data berada pada batas antara dua *cluster*.

III. METODE PENELITIAN

Penelitian terdahulu yang menggambarkan metode DBSCAN ialah pada penelitian yang dilakukan oleh (Tukiyat & Djohan, 2022) dengan judul “Analisis Penyebaran Pandemi Covid-19 Di Kota Jakarta Menggunakan Metode Clustering K-Means Dan *Density Based Spatial Clustering of Application With Noise* (DBSCAN)” yang dimana penelitian ini dilakukan untuk melakukan klasterisasi pada data Covid-19 didunia dengan menggunakan teknik *clustering* menggunakan algoritma DBSCAN dan juga *K-Means* dengan melakukan pertimbangan terhadap nilai validasi *cluster* terbaik sebagai penentu keakuratan algoritma yang digunakan. Pada pengolahan Dataset dengan metode K-Means dan DBSCAN ini

peneliti menggunakan bantuan dari aplikasi WEKA untuk menentukan jumlah *cluster* optimal dalam proses clustering ini dengan menggunakan metode *elbow*. Proses penentuan jumlah *cluster* terbaik menggunakan metode *Elbow* dengan data sampel berjumlah 151 data, sampel yang digunakan adalah urutan 20 sampai 171 dari dataset yang ada dengan uji coba jumlah *cluster* dari $K=3$ sampai $K=10$. Pada penelitian.

Pada judul “Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia” (Adha et al., 2021) pada penelitian ini proses klasterisasi dengan menggunakan algoritma DBSCAN dilakukan percobaan sebanyak 22 kali yang Dimana setiap percobaan menggunakan nilai *epsilon* (*Eps*) dan *MinPoints* (*MinPts*) yang berbeda. Nilai *Eps* yang digunakan ialah pada rentang 0,1 hingga 0,2. Sedangkan nilai *MinPts* ialah 3 dan 4. Kemudian dilakukan uji validasi pada *cluster* dengan menggunakan SI dengan hasil yang diperoleh 0,3624 dari percobaan dengan nilai *Eps* 0,2 dan nilai *MinPts* 3. Perbandingan algoritma DBSCAN dan *K-Means* dilakukan dengan menentukan jumlah *cluster* yang paling optimal dilihat dari nilai *Silhouette Index* (SI). klasterisasi menggunakan DBSCAN yang memiliki nilai SI terbaik yaitu 0.3624 dengan nilai *Eps* 0,2 dan nilai *MinPts* 3. Sementara dari hasil klasterisasi menggunakan *K-Means*, nilai SI terbaik diperoleh percobaan $k=8$ dengan nilai 0,6902. Maka pada penelitian ini, algoritma *K-Means* memiliki nilai validitas *cluster* lebih baik dibandingkan algoritma DBSCAN.

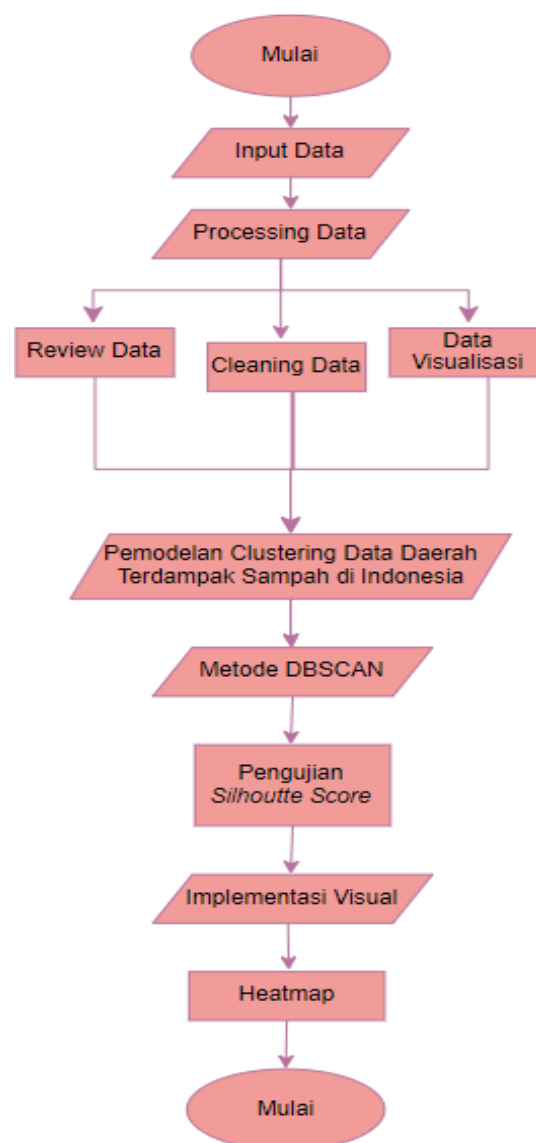
Sedangkan pada penelitian (Chen et al., 2020) dengan judul “Analisis Data Sebaran Penyakit Menggunakan Algoritma *Density Based Spatial Clustering Of Applications With Noise*” Pada penelitian tersebut dalam penentuan nilai *epsilon* dan nilai *Min Points* sangat mempengaruhi jumlah *cluster* yang

terbentuk. Pada penyakit ISPA persebaran yang tertinggi ialah pada *cluster* I dengan daerah Purwakarta dan Babakancikao dari 5 *cluster* yang ada. Kemudian untuk penyakit TBC persebaran *cluster* yang tertinggi ialah pada *cluster* 4 dari 4 *cluster* yang ada pada daerah Bungursari dan Campaka, lalu pada penyakit Asma persebaran *cluster* tertinggi ialah pada *cluster* 4 dari 4 *cluster* yang ada dengan daerah Bungursari dan Cempaka dengan *cluster* daerah tertinggi. Dan terakhir pada penyakit *Pneumonia* persebaran *cluster* yang tertinggi ialah pada *cluster* 4 dari 4 *cluster* yang ada ada dengan daerah Bungursari dan Cempaka sebagai daerah dengan *cluster* tertinggi.

Data yang digunakan dalam penelitian ini ialah data yang bersumber dari *website* resmi Kementerian Lingkungan Hidup dan Kehutanan (KLHK) pada Sistem Informasi Pengelolaan Sampah Nasional (SIPSN). Dimana data yang digunakan pada penelitian ini merupakan data dari tahun 2020 hingga 2022. Jenis penelitian ini menggunakan jenis desain penelitian kuantitatif dengan menggunakan metode data mining. Proses analisis data dilakukan melalui tahapan *preprocessing*, normalisasi (*Min-Max Scaling*), dan evaluasi menggunakan *Silhouette Score*, yang merupakan indikator kuantitatif kualitas *cluster*. Tujuan utama dari penelitian ini ialah menentukan pola persebaran daerah yang terdampak sampah berdasarkan karakteristik numerik pada data. Langkah awal yang dilakukan Dalam penelitian ini ialah studi literatur guna memahami lebih Dalam tentang penelitian yang ingin dilakukan. Selanjutnya peneliti membuat rumusan masalah dan tujuan yang ingin dicapai pada penelitian ini agar

penelitian dapat lebih terarah berdasarkan studi literati yang telah dilakukan oleh peneliti sebelumnya. Selanjut peneliti melakukan pengumpulan data. Yang Dimana Sebagian besar data yang peneliti gunakan bersumber dari *website* resmi Kementerian Lingkungan Hidup dan Kehutanan (KLHK) pada Sistem Informasi Pengelolaan Sampah Nasional (SIPSN). Dan pada Sebagian data dilakukan proses normalisasi (*Min-Max Scaling*) guna mengubah nilai suatu data agar dapat berada dalam rentang yang sama. Jika proses normalisasi data tidak dilakukan maka variabel dengan nilai lebih besar dapat mendominasi dalam perhitungan jarak dan dapat memengaruhi hasil clustering dari proses normalisasi tersebut, sehingga dapat menghasilkan nilai Indeks Dampak Sampah (IDS) yang digunakan dalam mengelompokkan daerah-daerah yang terdampak sampah di Indonesia.

Langkah selanjutnya ialah pengolahan data menggunakan metode DBSCAN dengan menggunakan *google colab*. Pada percobaan yang dilakukan menggunakan beberapa nilai parameter *Eps* dan *MinPts* hingga didapatkan nilai parameter yang optimal kemudian untuk mengukur nilai validitas *cluster* digunakan metode *Silhouette Score*. Dan hasil akhir dari penelitian ini berupa visualisasi *heatmap* guna menampilkan daerah-daerah yang terdampak sampah dengan dilakukan pembagian dari beberapa zona



Gambar 3.1 Metodologi Penelitian

IV. HASIL PENELITIAN DAN PEMBAHASAN

Tahapan yang dilakukan pada pembahasan ini antara lain tahapan pengumpulan data dan *preprocessing* data, penerapan algoritma DBSCAN, dan Visualisasi. Adapun pembahasannya ialah sebagai berikut ;

A. Pengumpulan Dan *Preprocessing* Data

Dalam penelitian ini bersumber dari *website* resmi Kementerian Lingkungan Hidup dan Kehutanan (KLHK) pada Sistem Informasi Pengelolaan Sampah Nasional (SIPSN). Data yang digunakan pada penelitian ini memiliki 22 atribut yang terdiri dari “Tahun, Provinsi, Kabupaten/Kota, *Latitude*, *Longitude*, Jumlah Penduduk, Luas Wilayah (km^2), Kepadatan Penduduk, Timbulan Sampah Harian(ton), Timbulan Sampah Tahunan(ton), Jumlah Kepadatan Timbulan Sampah Harian(kg/org/hari), Jumlah Kepadatan Timbulan Sampah Tahunan ($\text{km}^2/\text{org}/\text{tahun}$), Sampah Terkelola Tahunan (ton/tahun), Sampah Tidak Terkelola Tahunan (ton/tahun), Rasio Sampah Terkelola(%), Rasio Sampah Tidak Terkelola(%), Normalisasi Kepadatan Penduduk, Normalisasi Kepadatan Timbulan Sampah Tahunan($\text{km}^2/\text{org}/\text{tahun}$), Normalisasi Rasio Sampah Tidak Terkelola, dan Indeks Dampak Sampah”. yang Dimana pada penelitian ini dilakukan penelitian daerah terdampak sampah di Indonesia pada tahun 2020,2021 dan 2022.

Pada data yang digunakan terdapat beberapa data yang melalui tahapan normalisasi (*Min-Max Scaling*) seperti Kepadatan Penduduk, Timbulan Sampah Harian(ton), Timbulan Sampah Tahunan(ton) dan Rasio Sampah Tidak Terkelola(%). Hal tersebut) diterapkan dalam penelitian ini guna mengubah nilai suatu data agar dapat berada dalam rentang yang sama. Jika proses normalisasi data tidak dilakukan maka variabel dengan nilai lebih besar dapat mendominasi dalam perhitungan jarak dan dapat memengaruhi hasil *clustering* dari proses normalisasi tersebut, sehingga dapat menghasilkan nilai Indeks Dampak Sampah (IDS) yang digunakan dalam

mengelompokkan daerah-daerah yang terdampak sampah di Indonesia.

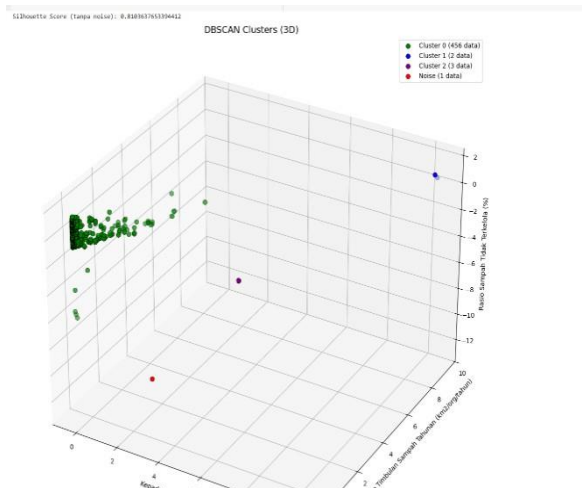
Tahun	Provinsi	Kabupaten/Kota	Normalisasi Kepadatan Penduduk	Normalisasi Kepadatan Timbulan Sampah Tahunan ($\text{km}^2/\text{org}/\text{tahun}$)	Normalisasi Rasio Sampah Tidak Terkelola
2021	Aceh	Kota Banda Aceh	0.08	0.24	0.69
2021	Aceh	Kota Sabang	0.01	0.01	0.70
2021	Aceh	Kota Lingsa	0.01	0.01	0.74
2021	Sumatera Utara	Kab. Langkat	0.00	0.00	0.96
2021	Sumatera Utara	Kab. Deli Serdang	0.01	0.01	0.85
2021	Sumatera Utara	Kota Medan	0.17	0.38	0.81
2021	Sumatera Utara	Kota Tebing Tinggi	0.08	0.18	0.70
2021	Sumatera Barat	Kab. Tanah Datar	0.01	0.01	0.84
2021	Sumatera Barat	Kab. Agam	0.00	0.01	0.92
2021	Sumatera Barat	Kab. Lima Puluh Kota	0.00	0.00	0.91
2021	Sumatera Barat	Kab. Dharmasraya	0.00	0.00	0.72
2021	Sumatera Barat	Kota Solok	0.05	0.05	0.88
2021	Sumatera Barat	Kota Sawahlunto	0.00	0.00	0.70
2021	Sumatera Barat	Kota Padang Panjang	0.05	0.12	0.68
2021	Sumatera Barat	Kota Payakumbuh	0.03	0.07	0.69

Gambar 4.1 Contoh Proses Normalisasi Data Pada Tahun 2022

B. *Clustering* Menggunakan Algoritma DBSCAN

Proses *Clustering* pada penelitian ini menggunakan algoritma DBSCAN yang menghasilkan *cluster* sebanyak 3 kali, yang dimana pengelompokkan data pada *cluster* utama dikelompokkan berdasarkan kemiripan pola datanya. Namun, pada data tersebut didapatkan 1 daerah yang dianggap sebagai *outlier* (data asing) oleh DBSCAN. Hal itu dikarekan data tersebut memiliki karakteristik yang berbeda dengan data lainnya. Dalam hal ini dapat terjadi ketika suatu daerah memiliki jumlah kepadatan penduduk atau jumlah kepadatan sampah yang sangat ekstrem dibandingkan daerah lain. Nilai *Eps* optimal yang digunakan dalam penelitian ini ialah 2 dan *MinPts* optimal juga 2. Hasil *cluster* dari data yang diteliti oleh peneliti menunjukkan hasil yang baik. Hal itu dapat dilihat dari hasil *Silhouette Score* (0.810). Adapun fitur yang digunakan dalam DBSCAN pada penelitian ini ialah Kepadatan Penduduk, Jumlah Kepadatan Timbulan Sampah Tahunan ($\text{km}^2/\text{org}/\text{tahun}$) dan Rasio Sampah

Terkelola (%). Hasil DBSCAN dapat dilihat pada **Gambar 4.2**



Gambar 4.2 Visualisasi *Clustering* Menggunakan Algoritma DBSCAN

Pada **Gambar 4.2** masing-masing *cluster* terbagi menjadi warna yang berbeda berdasarkan keterangan yang ada. Dan *cluster* terbanyak terletak pada *cluster 0* dengan 456 data.

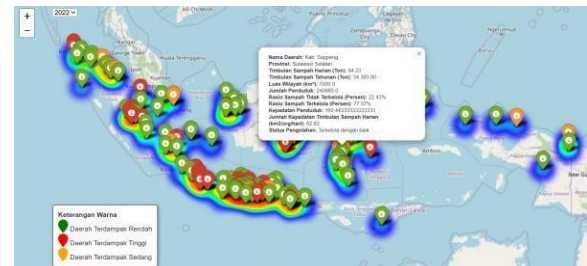
C. Visualisasi *Heatmap*



Gambar 4.3 *Heatmap* 2020



Gambar 4.4 *Heatmap* 2021



Gambar 4.5 *Heatmap* 2022

Pada **Gambar 4.3**, **Gambar 4.4** dan **Gambar 4.5** menampilkan *Heatmap* pada tahun 2020,2021 dan 2022. Gambar tersebut merupakan *map* interaktif yang dilengkapi dengan visualisasi *heatmap* guna untuk menunjukkan tingkat persebaran sampah di berbagai daerah di Indonesia. Yang dimana pada *heatmap* tersebut menunjukkan daerah yang terdampak sampah pada suatu daerah. Warna Merah menunjukkan wilayah dengan intensitas tinggi yang dimana daerah tersebut memproduksi sampah dalam jumlah besar atau memiliki permasalahan dalam pengelolaan sampahnya. Warna kuning menunjukkan wilayah dengan intensitas yang sedang, yang dimana Tingkat produksi sampahnya cukup signifikan namun tidak sebesar zona merah. Dan untuk zona biru merupakan wilayah dengan intensitas rendah. Yang dimana menunjukkan jumlah sampah yang lebih sedikit atau pengelolaan yang lebih baik.

V. KESIMPULAN

Jadi, dapat disimpulkan bahwa untuk mendapatkan hasil clusterisasi pada daerah terdampak sampah yang ada di Indonesia dibutuhkan beberapa tolak ukur suatu daerah dapat dikatan terdampak sampah dan tidak terdampak dapat dilihat dari nilai Indeks Dampak Sampah (IDS) yang dimana nilai IDS

didapatkan dari hasil penjumlahan nilai yang telah di normalisasi pada kepadatan penduduk, Kepadatan Timbulan Sampah Harian ($\text{km}^2/\text{org}/\text{hari}$) dan Rasio Sampah Tidak Terkelola.

Proses *Clustering* pada penelitian ini menggunakan algoritma DBSCAN yang menghasilkan *cluster* sebanyak 3 kali, yang dimana pengelompokan data pada *cluster* utama dikelompokkan berdasarkan kemiripan pola datanya. Namun, pada data tersebut didapatkan 2 daerah yang dianggap sebagai *outlier* (data asing) oleh DBSCAN. Hal itu dikarekan data tersebut memiliki karakteristik yang berbeda dengan data lainnya. Adapun hasil *cluster* dari data yang diteliti oleh peneliti menunjukkan hasil yang baik. Hal itu dapat dilihat dari hasil *Silhouette Score* (0.810).

DBSCAN untuk Clustering Seleksi Penentuan Mahasiswa yang Berhak Menerima Beasiswa Yayasan. 325–331.
Tukiyat, T., & Djohan, Y. (2022). Analisis Penyebaran Pandemi Covid-19 Di Kota Jakarta Menggunakan Metode Clustering K-Means Dan Density Based Spatial Clustering of Application With Noise (Dbscan). *Jurnal Informatika*, 9(1), 43–54.
<https://doi.org/10.31294/inf.v9i1.11226>

DAFTAR PUSTAKA

- Adha, R., Nurhaliza, N., Sholeha, U., & Mustakim, M. (2021). Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 18(2), 206–211.
- Asriati, N., & Harjanti, D. T. (2023). *PERSEPSI MASYARAKAT TENTANG PENGELOLAAN SAMPAH DI KOTA PONTIANAK* Wahyu Rachmadi, 2. 1(2), 44–57.
- Chen, H., Xu, H., Yenti Sumarni, Siaha Widodo, A., Pujayanti, D. A., Ishatono, I., Raharjo, S. T., Aristi, N. M., & Pratama, A. R. (2020). M 2 -1,2-. *Al Intaj: Jurnal Ekonomi Dan Perbankan Syariah*, 6(2), 159.
<http://jurnal.umt.ac.id/index.php/nyimak>
- Indini, D. P., Siburian, S. R., & Utomo, D. P. (2022). *Implementasi Algoritma*