

ANALISIS PENGELOLAAN SAMPAH DENGAN K-MEANS CLUSTERING DI INDONESIA

KELOMPOK 14

RENDI YULIO PRAMUDITA 2023071047
FAYYADH AHMAD MURHALI 2023071065

PERMASALAHAN YANG TERJADI DI MASYARAKAT

Indonesia menghasilkan lebih dari 40 juta ton sampah per tahun, dengan sebagian besar berasal dari limbah rumah tangga dan plastik. Tantangan dalam pengelolaan sampah mencakup pengumpulan, pemilahan, dan daur ulang yang belum optimal. Salah satu masalah utama dalam pengelolaan sampah di Indonesia adalah distribusi sumber daya yang tidak merata. Dengan pengelompokan daerah berdasarkan data produksi sampah dan kapasitas pengelolaannya, alokasi sumber daya dapat dioptimalkan.

ALGORITMA K-MEANS

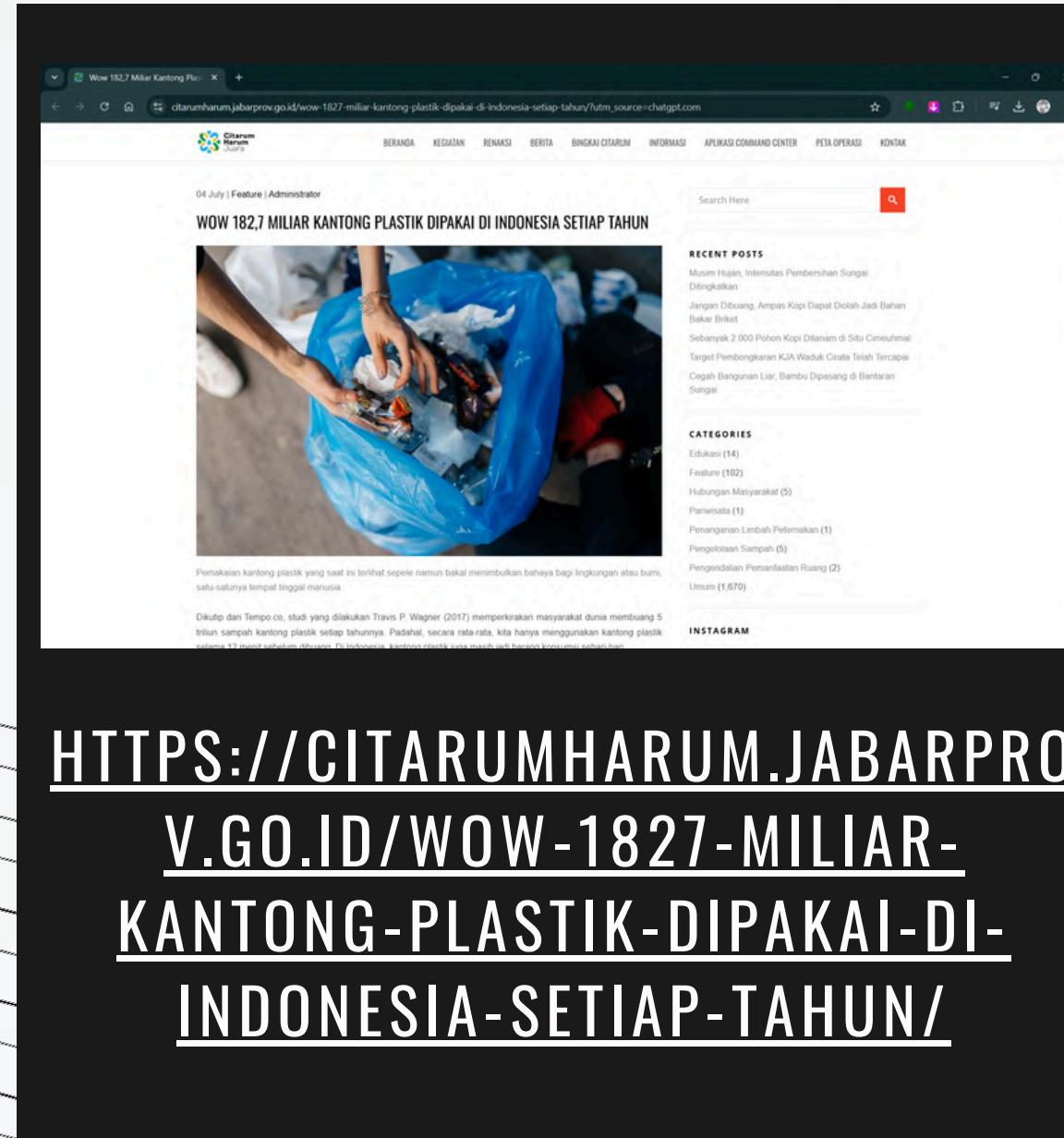
Algoritma K-Means adalah metode clustering yang membagi data ke dalam k kelompok berdasarkan kedekatan antar data. Proses dimulai dengan memilih k centroid acak, lalu data dikelompokkan berdasarkan jarak terdekat ke centroid. Proses dimulai dengan memilih k centroid acak, lalu data dikelompokkan berdasarkan jarak terdekat ke centroid. Posisi centroid diperbarui dengan rata-rata posisi data dalam kelompok tersebut, dan proses ini diulang hingga centroid tidak berubah. K-Means sederhana, cepat, dan cocok untuk pola kelompok yang jelas, tetapi sensitif terhadap outlier dan memerlukan penentuan jumlah kelompok k . Dalam penanggulangan sampah, K-Means bisa digunakan untuk mengelompokkan wilayah dan menentukan lokasi fasilitas pengelolaan sampah yang efektif.

BERITA PENDUKUNG



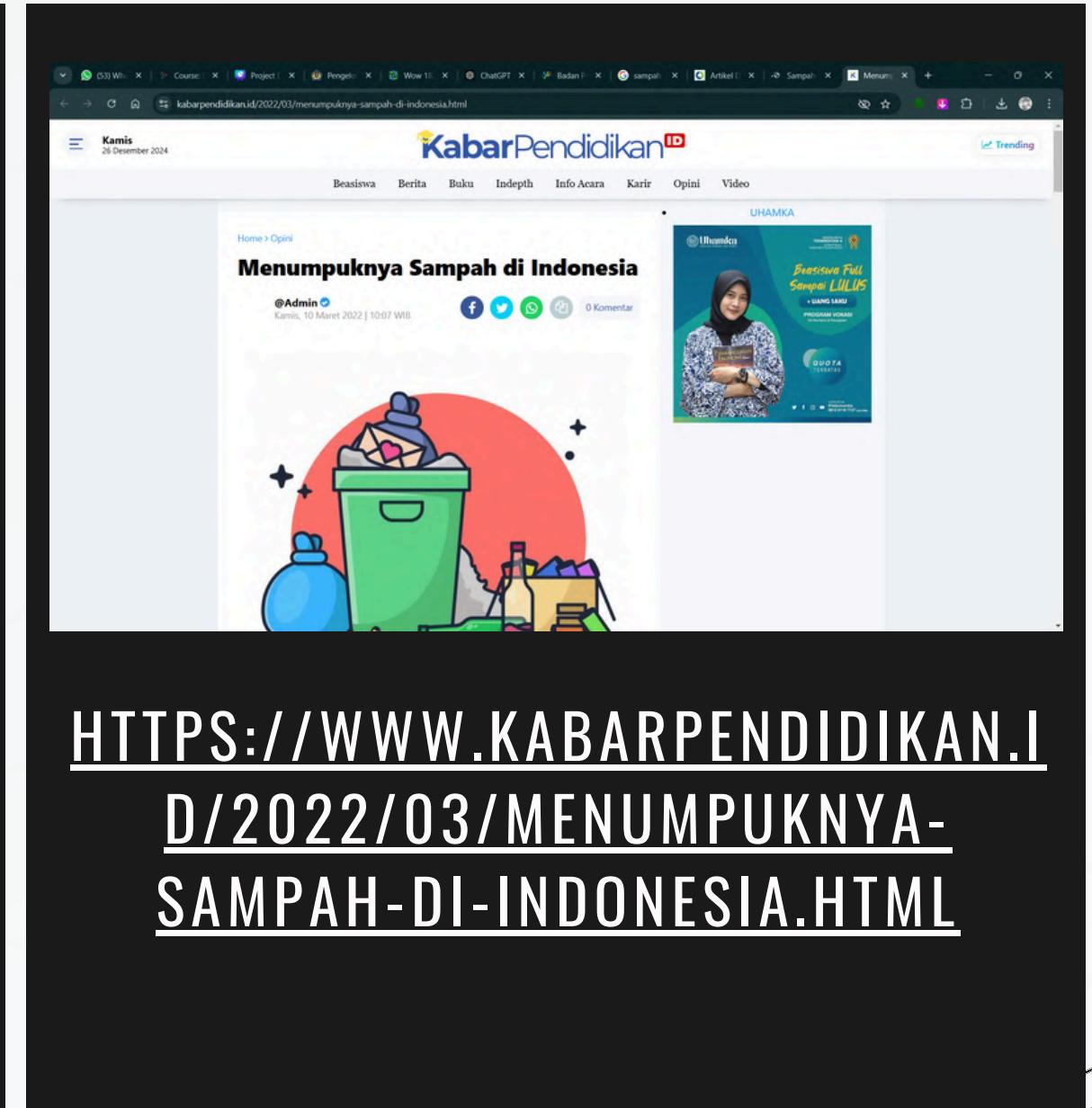
<https://www.djkn.kemenkeu.go.id/kpnl-lahat/baca-artikel/14891/pengelolaan-sampah-di-indonesia.html>

The screenshot shows a news article from the Ministry of Finance (Kemenkeu) titled "Pengelolaan Sampah Rumah Tangga dan Sampah Sejenis Sampah Rumah Tangga". It features a diagram illustrating the recycling process: household waste and similar household waste are sorted into four bins (Glass, Paper, Plastic, Bottles), which are then sent to a TPS (Treatment Plant). A small icon of a person is also present.



<https://citarumharum.jabarprov.go.id/wow-1827-miliar-kantong-plastik-dipakai-di-indonesia-setiap-tahun/>

The screenshot shows a news article from Citarum Harum titled "Wow 182.7 Miliar Kantong Plastik Dipakai di Indonesia Setiap Tahun". It includes a photograph of hands putting trash into a blue plastic bag. The sidebar contains a search bar, recent posts, categories, and an Instagram section.



<https://www.kabarpendidikan.id/2022/03/menumpuknya-sampah-di-indonesia.html>

The screenshot shows a news article from KabarPendidikan titled "Menumpuknya Sampah di Indonesia". It features a large illustration of a green trash bin overflowing with trash against a red background. A small photo of a woman is visible on the right side of the page.

STUDI PUSTAKA

Beberapa penelitian sebelumnya telah menunjukkan keberhasilan algoritma K-Means dalam menyelesaikan masalah clustering:

- 1.Jain A, K. (2010): Penerapan K-Means dalam pengelompokan pelanggan di bidang ritel.
- 2.Zhao W. (2015): Penggunaan K-Means untuk analisis pola lalu lintas kota.
- 3.Prasetyo H. (2020): Penerapan K-Means untuk optimasi pengelolaan limbah industri di Indonesia.





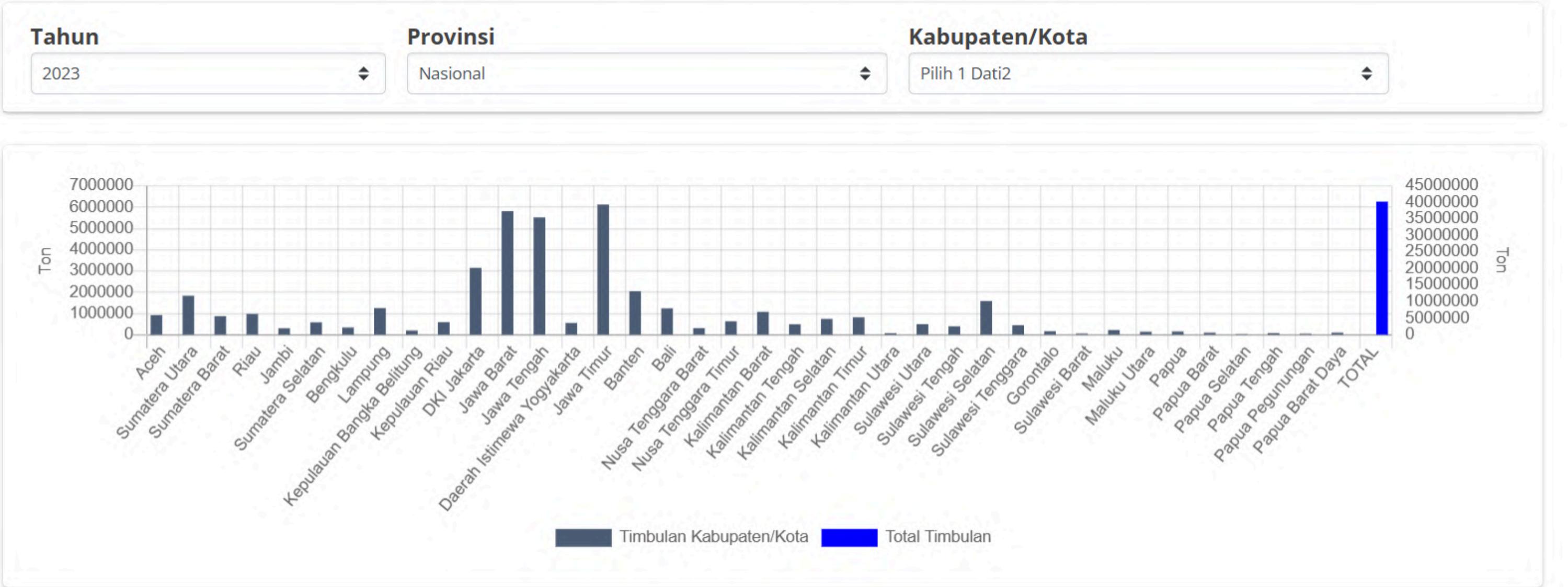
RELATED WORK

- 1. Pengelolaan Sumber Daya Energi:**
Studi oleh Lee S. (2018) menggunakan K-Means untuk mengelompokkan daerah berdasarkan konsumsi energi dan efisiensi penggunaan.
- 2. Pengelolaan Sampah Perkotaan:**
Gupta R. (2019) menggunakan K-Means untuk mengelompokkan zona perkotaan berdasarkan kapasitas tempat pembuangan sampah.
- 3. Analisis Sebaran Fasilitas Pelayanan Kesehatan:**
Chen, X. (2020). Analisis Sebaran Fasilitas Pelayanan Kesehatan Menggunakan Teknik Clustering

DATA SAMPAH

Data Dari Sistem Informasi Pengelolaan Sampah Nasional: SIPSN

TIMBULAN SAMPAH



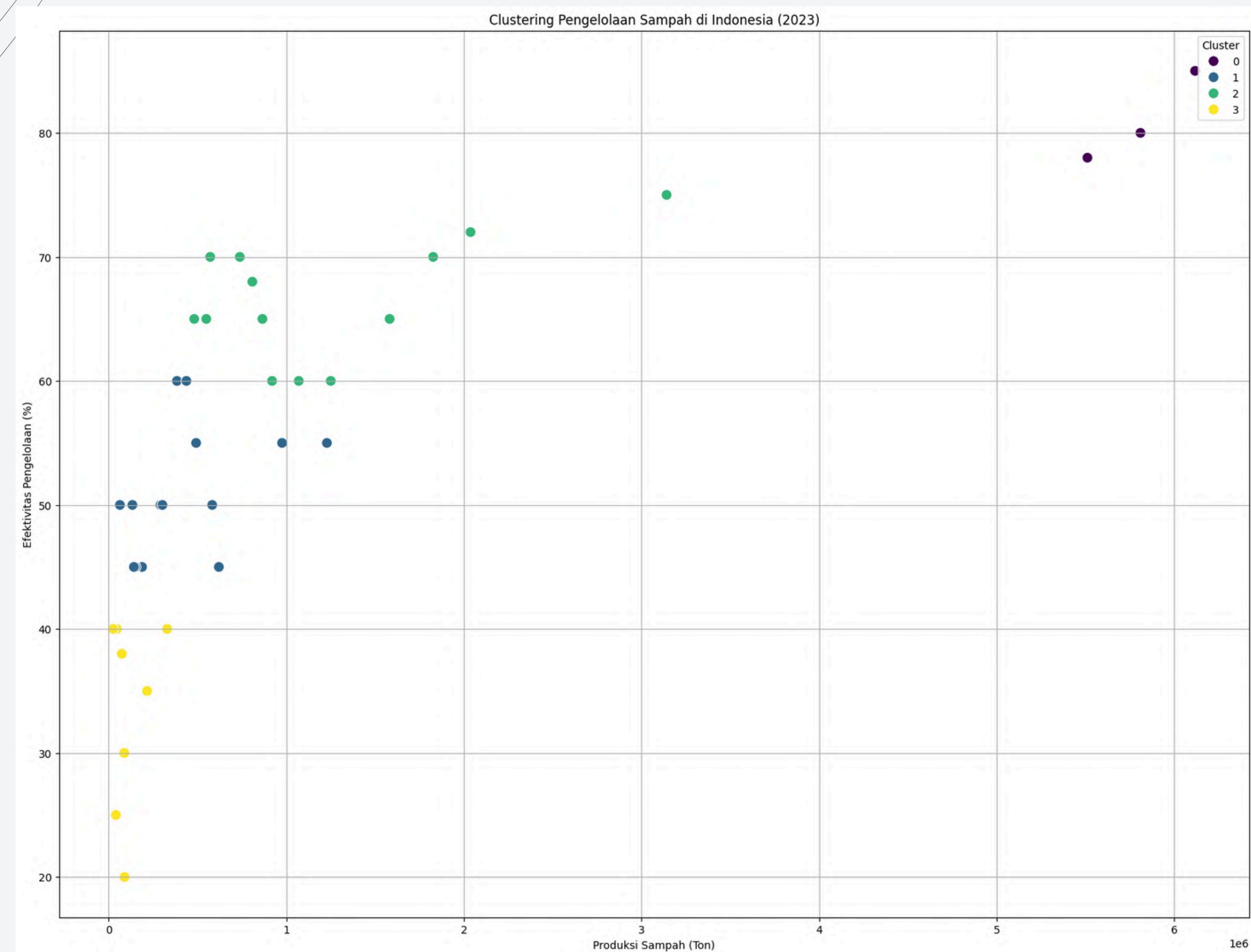
CODE PYTHON

```
UAS BasDat.py X
UAS BasDat.py > ...
1 import pandas as pd
2 import numpy as np
3 from sklearn.cluster import KMeans
4 import matplotlib.pyplot as plt
5 import seaborn as sns
6
7 # Membaca data dari tiga file Excel
8 produksi_file = "Produksi_Sampah.xlsx"
9 efektivitas_file = "Efektivitas_Pengelolaan.xlsx"
10 provinsi_file = "Provinsi.xlsx"
11
12 # Membaca file Excel
13 df_produksi = pd.read_excel(produksi_file)
14 df_efektivitas = pd.read_excel(efektivitas_file)
15 df_provinsi = pd.read_excel(provinsi_file)
16
17 # Validasi keberadaan kolom yang diperlukan
18 if "Produksi_Sampah" not in df_produksi.columns:
19     raise KeyError("Kolom 'Produksi_Sampah' tidak ditemukan di file Produksi_Sampah.xlsx")
20 if "Efektivitas_Pengelolaan" not in df_efektivitas.columns:
21     raise KeyError("Kolom 'Efektivitas_Pengelolaan' tidak ditemukan di file Efektivitas_Pengelolaan.xlsx")
22 if "Provinsi" not in df_provinsi.columns:
23     raise KeyError("Kolom 'Provinsi' tidak ditemukan di file Provinsi.xlsx")
24
25 # Menggabungkan ketiga dataframe berdasarkan indeks
26 df = pd.concat([df_provinsi[["Provinsi"]], df_produksi[["Produksi_Sampah"]], df_efektivitas[["Efektivitas_Pengelolaan"]]], axis=1)
27
28 # Normalisasi data (penting untuk clustering)
29 features = ["Produksi_Sampah", "Efektivitas_Pengelolaan"]
30 df_scaled = (df[features] - df[features].mean()) / df[features].std()
31
32 # Inisialisasi dan penerapan K-Means
33 kmeans = KMeans(n_clusters=4, random_state=42)
34 df["Cluster"] = kmeans.fit_predict(df_scaled)
35
36 # Visualisasi hasil clustering
37 plt.figure(figsize=(60, 50))
38 sns.scatterplot(
39     x="Produksi_Sampah",
40     y="Efektivitas_Pengelolaan",
41     hue="Cluster",
42     data=df,
43     palette="viridis",
44     s=100
45 )
46
47 plt.title("Clustering Pengelolaan Sampah di Indonesia (2023)")
48 plt.xlabel("Produksi Sampah (Ton)")
49 plt.ylabel("Efektivitas Pengelolaan (%)")
50 plt.legend(title="Cluster")
51 plt.grid(True)
52 plt.show()
53
54 # Output hasil clustering
55 print(df)
56
57
```

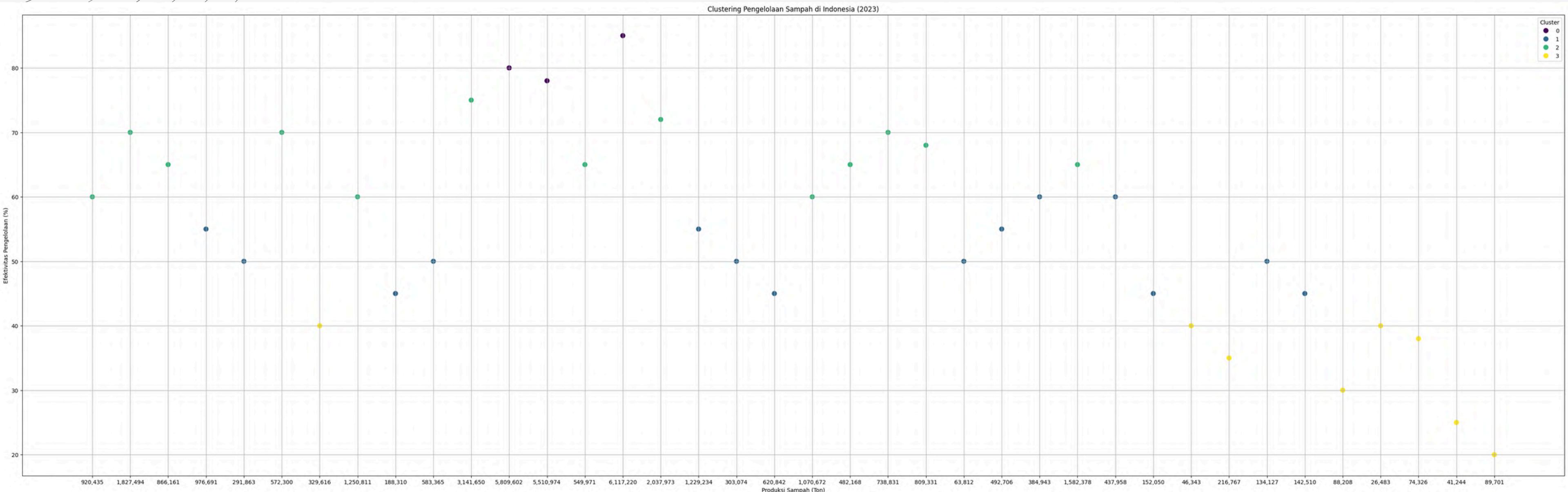
HASIL

	Provinsi	Produksi_Sampah	Efektivitas_Pengelolaan	Cluster
0	Aceh	920435	60	2
1	Sumatera Utara	1827494	70	2
2	Sumatera Barat	866161	65	2
3	Riau	976691	55	1
4	Jambi	291863	50	1
5	Sumatera Selatan	572300	70	2
6	Bengkulu	329616	40	3
7	Lampung	1250811	60	2
8	Kepulauan Bangka Belitung	188310	45	1
9	Kepulauan Riau	583365	50	1
10	DKI Jakarta	3141650	75	2
11	Jawa Barat	5809602	80	0
12	Jawa Tengah	5510974	78	0
13	Daerah Istimewa Yogyakarta	549971	65	2
14	Jawa Timur	6117220	85	0
15	Banten	2037973	72	2
16	Bali	1229234	55	1
17	Nusa Tenggara Barat	303074	50	1
18	Nusa Tenggara Timur	620842	45	1
19	Kalimantan Barat	1070672	60	2
20	Kalimantan Tengah	482168	65	2
21	Kalimantan Selatan	738831	70	2
22	Kalimantan Timur	809331	68	2
23	Kalimantan Utara	63812	50	1
24	Sulawesi Utara	492706	55	1
25	Sulawesi Tengah	384943	60	1
26	Sulawesi Selatan	1582378	65	2
27	Sulawesi Tenggara	437958	60	1
28	Gorontalo	152050	45	1
29	Selawesi Barat	46343	40	3
30	Maluku	216767	35	3
31	Maluku Utara	134127	50	1
32	Papua	142510	45	1
33	Papua Barat	88208	30	3
34	Papua Selatan	26483	40	3
35	Papua Tengah	74326	38	3
36	Papua Pegunungan	41244	25	3
37	Papua Barat Daya	89701	20	3

HASIL



HASIL



HASIL ANALISIS

Algoritma K-Means mengelompokkan provinsi di Indonesia ke dalam empat cluster berdasarkan produksi sampah (ton) dan efektivitas pengelolaan (persen). Cluster 0 mencakup provinsi dengan produksi sampah tinggi dan efektivitas pengelolaan baik, seperti DKI Jakarta dan Jawa Timur. Cluster 1 berisi provinsi dengan produksi sedang dan efektivitas tinggi, seperti Bali dan Yogyakarta. Cluster 2 mencakup provinsi dengan produksi rendah dan efektivitas rendah, seperti Papua Barat Daya. Cluster 3 terdiri dari provinsi dengan produksi rendah hingga sedang dan efektivitas sedang, seperti Sulawesi Tengah. Scatter plot menunjukkan distribusi antar cluster, dengan Jawa Timur memiliki produksi sampah tertinggi (6.117.220 ton) dan Jawa Barat efektivitas pengelolaan terbaik (80%).

KESIMPULAN

Kesimpulan



Hasil clustering menunjukkan hubungan antara produksi sampah dan efektivitas pengelolaan di setiap provinsi, yang dapat digunakan untuk menentukan prioritas pengelolaan sampah. Provinsi dengan produksi sampah tinggi dan efektivitas rendah (Cluster 2) perlu mendapat alokasi sumber daya lebih melalui peningkatan fasilitas, teknologi, dan kampanye kesadaran. Daerah di Cluster 0 perlu fokus pada peningkatan efisiensi dengan teknologi modern, sementara Cluster 2 dan 3 memerlukan edukasi tentang pengelolaan sampah. Temuan ini membantu pengambil kebijakan menyusun program berbasis data, seperti alokasi anggaran dan kampanye lingkungan.

TERIMA KASIH

Untuk data: <https://github.com/RendiYulio/DAA>

