
CLUSTERING KUALITAS UDARA MENGGUNAKAN ALGORITMA K-MEANS DI KOTA PONTIANAKMunawar¹, Putri Yuli Utami², Rizki Surtiyan Surya³^{1,2,3}Universitas Muhammadiyah Pontianakmunawarjos6@gmail.com¹, putriyuli@unmuhpnk.ac.id²,
rizki.surtiyan@unmuhpnk.ac.id³**ABSTRAK**

Pontianak sebagai ibu kota Provinsi Kalimantan Barat tengah menghadapi tantangan besar terkait peningkatan polusi udara akibat pesatnya urbanisasi dan industrialisasi. Pertumbuhan ekonomi yang cepat di kota ini tidak hanya membawa manfaat, tetapi juga menimbulkan dampak negatif terhadap lingkungan, terutama dalam bentuk peningkatan emisi polutan udara. Penelitian ini bertujuan untuk mengelompokkan data Indeks Standar Pencemar Udara (ISPU) di kota Pontianak berdasarkan konsentrasi polutan menggunakan algoritma *K-means*. Parameter ISPU yang dianalisis mencakup polutan PM₁₀, PM_{2.5}, SO₂, CO, O₃, NO₂ dan HC yang diperoleh dari tahun 2021 hingga 2023. *Elbow Method* digunakan untuk menentukan jumlah *cluster* yang optimal dan menghasilkan tiga *cluster* (*k*3). Evaluasi kualitas *cluster* dilakukan menggunakan *Davies Bouldin Index* (DBI) yang menunjukkan nilai DBI sebesar 0.881 untuk tahun 2021, 1.036 untuk tahun 2022, dan 1.011 untuk tahun 2023. Hasil *clustering* menunjukkan bahwa polutan PM_{2.5} memiliki konsentrasi tertinggi pada tahun 2021 dengan nilai 429 sedangkan PM₁₀ menjadi dominan pada tahun 2022 dan 2023 dengan nilai masing-masing 126 dan 154. Hasil penelitian ini memberikan informasi terkait pengambilan kebijakan yang tepat sebagai acuan pengendalian polusi udara yang lebih efektif, serta langkah-langkah dan antisipasi yang harus dilakukan oleh masyarakat.

Kata Kunci: *Clustering, Elbow Method, DBI, K-Means, Kualitas Udara.*

ABSTRACT

Pontianak as the capital of West Kalimantan Province is facing major challenges related to increasing air pollution due to rapid urbanization and industrialization. Rapid economic growth in this city not only brings benefits, but also has negative impacts on the environment, especially in the form of increased emissions of air pollutants. This research aims to group Air Pollution Standard Index (ISPU) data in Pontianak City based on pollutant concentrations using the K-means algorithm. The ISPU data analyzed includes PM₁₀, PM_{2.5}, SO₂, CO, O₃, NO₂ and HC pollutants obtained from 2021 to 2023. The Elbow method was used to determine the optimal number of clusters, resulting in three clusters (k=3). Cluster quality evaluation was carried out using the Davies Bouldin Index

(DBI) which shows a DBI value of 0.881 for 2021, 1.036 for 2022, and 1.011 for 2023. The clustering results show that $PM_{2.5}$ pollutants have the highest concentration in 2021 with a value of 429 while PM_{10} becomes dominant in 2022 and 2023 with values of 126 and 154 respectively. The results of this research provide information regarding appropriate policy making as a reference for more effective air pollution control, as well as steps and anticipation that must be taken by the community.

Keywords: *Clustering, Elbow Method, DBI, K-Means, Air Quality.*

A. PENDAHULUAN

Kualitas udara yang bersih menjadi bagian hal penting untuk keberlangsungan hidup dikarenakan udaralah yang membantu manusia bernafas, kondisi ini tentu saja berguna dan bermanfaat bagi tubuh manusia. Dampak negatif dari penurunan kualitas udara dapat berakibat buruk terhadap kesehatan manusia dan merusak lingkungan.

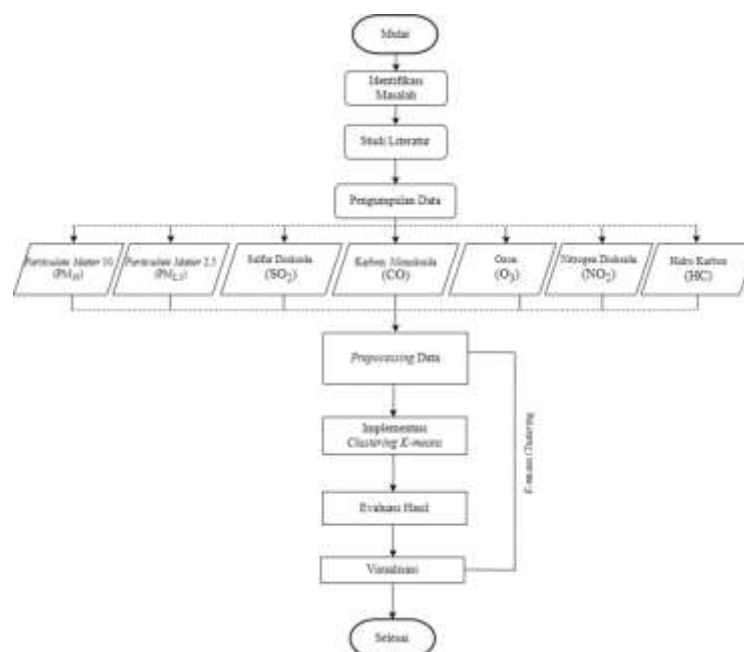
Dinas Lingkungan Hidup dan Kehutanan Provinsi Kalimantan Barat menyediakan data Indeks Standar Pencemar Udara untuk wilayah Kota Pontianak. Data ini dapat digunakan untuk memetakan dan memantau konsentrasi gas pencemar udara seperti Particulate Matter (PM_{10}), Particulate Matter ($PM_{2.5}$), Sulfur Dioksida (SO_2), Karbon Monoksida (CO), Ozon (O_3), Nitrogen Dioksida (NO_2), Hidrokarbon (HC). Berdasarkan Permen LHK No. 14 Tahun 2020 tentang Indeks Standar Pencemar Udara memiliki kategori yaitu rentang 0-50 menunjukkan kualitas udara baik, 51-100 menandakan kualitas udara sedang, 101-200 menunjukkan kualitas udara tidak sehat yang berpotensi merugikan manusia, hewan, dan tumbuhan. Rentang 201-300 menunjukkan kualitas udara sangat tidak sehat yang dapat meningkatkan risiko kesehatan pada kelompok sensitif. Di atas 300 menunjukkan kualitas udara berbahaya yang dapat secara serius merugikan kesehatan dan memerlukan tindakan cepat.

Alogaritma K-means memiliki kapasitas untuk mengelompokkan data dalam jumlah besar dengan waktu komputasi yang cepat dan efisien[3]. Clustering kualitas udara di Kota Pontianak menggunakan data Indeks Standar Pencemar Udara perlu dilakukan untuk mengelompokkan berbagai parameter yang mengandung zat-zat kimia berbahaya yang berasal dari aktivitas internal dan eksternal yang menyebabkan Pencemar udara. Hasil Clustering ini dapat menjadi informasi bagi pemerintah daerah, terutama dalam pengambilan kebijakan dan langkah-langkah yang tepat untuk mencegah serta mengatasi Pencemar udara.

Melihat latar belakang masalah di atas penulis tertarik melakukan penelitian dengan judul “Clustering kualitas udara menggunakan Algoritma K-means di Kota Pontianak”. Harapannya, hasil penelitian ini dapat meningkatkan kapasitas dan kesadaran manusia dan lembaga terkait dampak lingkungan, Meningkatkan efisiensi dan efektivitas program pengendalian Pencemar udara dan Meningkatkan kesadaran masyarakat tentang kualitas udara dan dampaknya terhadap kesehatan.

B. METODE PENELITIAN

Berikut digram alir dan penjelasan pada tahapan penelitian ini.



Gambar 1. Tahapan Penelitian

Tahapan penelitian dimulai dengan indentifikasi masalah, studi literatur, pengumpulan data, *preprocessing* data, implementasi *clustering K-means*, evaluasi hasil dan visualisasi.

Identifikasi Masalah

Permasalahan dalam penelitian ini diidentifikasi melalui studi pustaka dan literatur mengenai kualitas udara serta faktor pencemar udara di Kota Pontianak. Berdasarkan hasil identifikasi, rumusan masalah dalam penelitian ini meliputi: (1) bagaimana penerapan algoritma K-Means terhadap data Indeks Standar Pencemar Udara (ISPU) di

Kota Pontianak, (2) bagaimana kinerja *K-Means* berdasarkan nilai *Davies Bouldin Index* (DBI), dan (3) bagaimana visualisasi hasil *clustering* agar mudah dipahami.

Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data Indeks Pencemar Udara (ISPU) yang di peroleh dari Dinas Lingkungan Hidup dan Kehutanan Provinsi Kalimantan Barat dengan periode waktu 3 tahun terakhir yaitu tahun 2021-2023. Dataset yang diambil berupa PM₁₀, PM_{2.5}, SO₂, CO, O₃, NO₂, HC.

K-Means Clustering

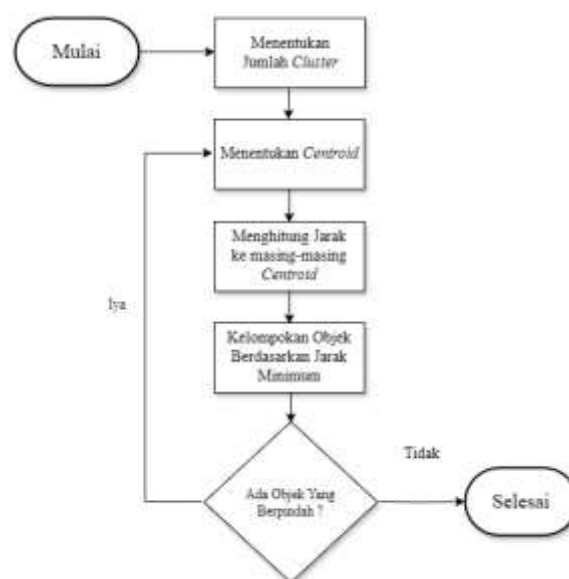
Pada tahap ini, dilakukan *preprocessing data*, implementasi *clustering K-means*, evaluasi *cluster* dengan DBI dan visualisasi.

a) *Preprocessing Data*

Tujuan utama dari proses ini adalah untuk meningkatkan kualitas hasil pengelompokan data saat menggunakan algoritma *K-means*[5]. *Preprocessing* data meliputi tahapan seleksi atribut. Pembersihan data (*Data Cleaning*), Transformasi data.[6]

b) *Implentasi Clustering K-Means*

Implementasi algoritma *K-means Clustering* dengan menggunakan rumus yang telah ditetapkan.



Gambar 2. Implentasi *Clustering K-Means*

c) Evaluasi hasil *K-Means* dengan DBI

Selanjutnya menguji performa pengelompokan dengan menggunakan metode validitas DBI (*Davies Bouldin Index*) Click or tap here to enter text..



Gambar 3. Evaluasi hasil K-Means dengan DBI

d) Visualisasi

Tujuan visualisasi ini adalah untuk menyajikan informasi yang kompleks dengan cara yang intuitif dan memudahkan pengambilan keputusan.

C. HASIL DAN PEMBAHASAN

Hasil Pengumpulan Data

Dari data yang di peroleh terdapat atribut yaitu Tanggal, PM₁₀, PM_{2.5}, SO₂, CO, O₃, NO₂, HC, *MAX Component*, *Critical Component* dan Kategori. Data periode tahun 2021-2023.

Tabel 1. Atribut Data dan Penjelasanya

No	Nama Atribut	Satuan	Penjelasan
1	Tanggal	dd/mm/yyyy	Tanggal pengukuran kualitas udara
2	PM ₁₀	ug/m ³	<i>Particulate Matter</i> salah satu parameter yang diukur

3	PM _{2.5}	ug/m ³	Particulate Matter salah satu parameter yang diukur
4	SO ₂	ug/m ³	Sulful dioksida salah satu parameter yang diukur
5	CO	ug/m ³	Karbon monoksida salah satu parameter yang diukur
6	O ₃	ug/m ³	Ozon salah satu parameter yang diukur
7	NO ₂	ug/m ³	Nitrogen dioksida salah satu parameter yang diukur
8	HC	ppm	Hidrokarbon salah satu parameter yang diukur
9	MAX Component	Nilai yang tertinggi	Nilai tertinggi diantara seluruh jenis parameter yang diukur
10	Critical Component	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC	Parameter dengan hasil tertinggi diantara jenis seluruh parameter yang diukur
11	Kategori	Baik/Sedang/Tidak Sehat/Sangat Tidak Sehat/Berbahaya	Indeks Standar Pencemar Udara (ISPU) yang diukur dengan nilai baik, sedang, tidak sehat, sangat tidak sehat, berbahaya

Berikut pada tabel 4.2 data ISPU di Kota Pontianak tahun 2021 sampai 2023.

Tabel 2. Data ISPU Tahun 2021

Tanggal	PM ₁₀	PM _{2.5}	SO ₂	CO	O ₃	NO ₂	HC	MAX	Critical Component	Kategori
01/01/2021	15	42	0	35	0	22	0	42	PM _{2.5}	BAIK
02/01/2021	12	33	0	35	0	16	0	35	CO	BAIK
03/01/2021	9	26	0	36	0	19	0	36	CO	BAIK
04/01/2021	11	38	0	38	0	23	0	38	CO	BAIK
05/01/2021	11	40	0	38	0	24	75	75	HC	SEDANG
06/01/2021	18	55	0	39	0	23	78	78	HC	SEDANG
07/01/2021	12	44	0	38	0	23	78	78	HC	SEDANG

08/01/2021	12	43	0	36	0	24	86	86	HC	SEDAN G
09/01/2021	19	56	0	37	0	23	77	77	HC	SEDAN G
10/01/2021	16	53	0	36	0	24	76	76	HC	SEDAN G
11/01/2021	7	26	0	36	0	24	74	74	HC	SEDAN G
12/01/2021	9	33	0	36	0	20	45	45	HC	BAIK
13/01/2021	5	21	0	37	0	14	11	37	CO	BAIK
14/01/2021	7	20	0	38	0	14	11	38	CO	BAIK
-----	---	---	----	---	-	--	---	---	---	-----
29/12/2021	28	65	0	48	0	20	0	65	PM _{2.5}	SEDAN G
30/12/2021	54	105	0	49	0	21	0	105	PM _{2.5}	TIDAK SEHAT
31/12/2021	32	70	0	49	0	22	0	70	PM _{2.5}	SEDAN G

Tabel 3. Data ISPU Tahun 2022

Tanggal	PM ₁₀	PM _{2.5}	SO ₂	CO	O ₃	NO ₂	HC	MAK	Critical Component	Kategori
01/01/2022	49	93	0	49	4	22	0	93	PM _{2.5}	SEDAN G
02/01/2022	32	70	0	49	5	23	0	70	PM _{2.5}	SEDAN G
03/01/2022	14	44	0	48	5	23	0	48	CO	BAIK
04/01/2022	19	54	0	49	5	24	75	75	HC	SEDAN G
05/01/2022	20	55	0	49	7	24	69	69	HC	SEDAN G

06/01/202	19	55	0	49	8	24	54	55	PM _{2.5}	SEDAN
2										G
07/01/202	19	54	0	49	9	24	59	59	PM _{2.5}	SEDAN
2										G
08/01/202	34	74	0	49	10	25	70	74	PM _{2.5}	SEDAN
2										G
09/01/202	28	67	0	49	11	25	62	67	PM _{2.5}	SEDAN
2										G
10/01/202	18	53	0	49	12	25	64	64	HC	SEDAN
2										G
11/01/202	25	62	0	49	13	25	68	68	PM _{2.5}	SEDAN
2										G
12/01/202	16	49	0	49	14	25	65	65	HC	SEDAN
2										G
-----	---	---	----	---	-	--	---	---	---	-----
-----	---	---	----	---	-	--	---	---	---	-----
-----	---	---	----	---	-	--	---	---	---	-----
29/12/202	21	50	14	14	7	3	1	50	PM _{2.5}	BAIK
2										
30/12/202	22	51	16	14	7	3	1	51	PM _{2.5}	SEDAN
2										G
31/12/202	21	52	16	14	7	4	1	52	PM _{2.5}	SEDAN
2										G

Tabel 4. Data ISPU Tahun 2023

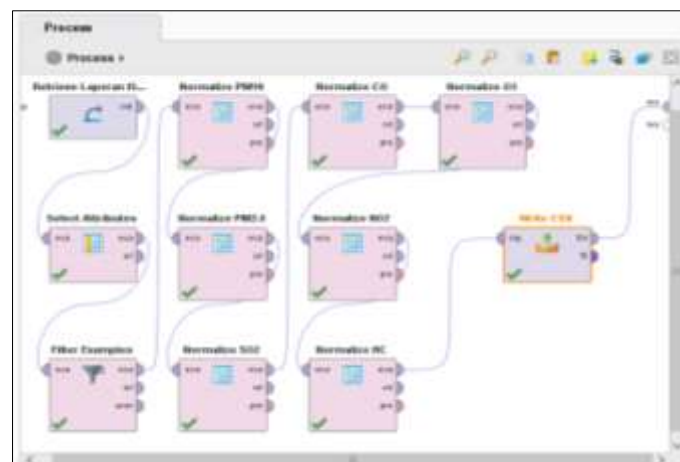
Tanggal	PM ₁₀	PM _{2.5}	SO ₂	CO	O ₃	NO ₂	HC	MA _X	Critical Component	Kategori
01/01/2023	25	56	17	14	7	4	1	56	PM _{2.5}	SEDAN G
02/01/2023	20	49	17	14	7	4	1	49	PM _{2.5}	BAIK
03/01/2023	25	56	16	14	7	3	1	56	PM _{2.5}	SEDAN G
04/01/2023	31	63	16	14	7	3	1	63	PM _{2.5}	SEDAN G
05/01/2023	30	61	23	14	7	4	1	61	PM _{2.5}	SEDAN G
06/01/2023	18	45	23	14	13	5	1	45	PM _{2.5}	BAIK
07/01/2023	25	57	16	14	13	4	1	57	PM _{2.5}	SEDAN G
08/01/2023	24	55	16	14	7	3	1	55	PM _{2.5}	SEDAN G
09/01/2023	22	53	17	14	7	2	1	53	PM _{2.5}	SEDAN G
10/01/2023	21	53	16	14	6	2	1	53	PM _{2.5}	SEDAN G
11/01/2023	38	72	15	14	7	2	1	72	PM _{2.5}	SEDAN G

12/01/202	30	63	35	14	7	4	1	63	PM _{2.5}	SEDAN
3										G
13/01/202	31	64	36	14	9	4	1	64	PM _{2.5}	SEDAN
3										G
14/01/202	22	54	14	14	9	4	1	54	PM _{2.5}	SEDAN
3										G
-----	---	---	---	---	-	--	---	---	---	-----
29/12/202	12	40	23	11	10	19	0	40	PM _{2.5}	BAIK
3										
30/12/202	8	27	23	11	10	19	0	27	PM _{2.5}	BAIK
3										
31/12/202	10	31	24	10	10	18	0	31	PM _{2.5}	BAIK
3										

K-Means Clustering

a) Preprocessing Data

preprocessing data yaitu menyiapkan data mentah sebelum digunakan dalam analisis atau model *machine learning*[9]. Proses ini untuk memastikan bahwa data tersebut berada dalam kondisi yang baik, bersih, dan siap untuk analisis lebih lanjut



Gambar 4. Tahap *Preprocessing* Data

Tahap ini melibatkan operator seperti operator select attributes, operator filter example, normalize dan operator write CSV. Berikut hasilnya.

Row No.	PM10	PM2.5	SO2	CO	O3	NO2	HC
1	20.663	52.825	12	37.909	1	24.959	18
2	17.730	44.077	12	37.909	1	19.571	18
3	14.798	37.273	12	38.727	1	22.265	18
4	16.753	48.937	12	40.364	1	25.857	18
5	16.753	50.881	12	40.364	1	26.755	81.362
6	23.596	65.462	12	41.182	1	25.857	83.897
7	17.730	54.769	12	40.364	1	25.857	83.897
8	17.730	53.797	12	38.727	1	26.755	90.655
9	24.573	65.434	12	39.545	1	25.857	83.052
10	21.640	63.517	12	38.727	1	26.755	82.207
11	12.843	37.273	12	38.727	1	26.755	80.517
12	14.798	44.077	12	38.727	1	23.163	56.017
13	10.888	32.413	12	39.545	1	17.776	27.293
14	12.843	31.441	12	40.364	1	17.776	27.293
15	14.798	35.329	12	40.364	1	16.878	23.914

ExampleSet (341 examples, 0 special attributes, 7 regular attributes)

Gambar 5. Hasil *Preprocessing* Data ISPU 2021

Row No.	PM10	PM2.5	CO	O3	SO2	NO2	HC
1	58.167	103.440	49	10.588	16	22	1
2	43.190	82.062	49	11.485	16	23	1
3	27.333	57.896	48	11.485	16	23	1
4	31.738	67.191	49	11.485	16	24	75.074
5	32.619	68.120	49	13.279	16	24	69.148
6	31.738	68.120	49	14.176	16	24	54.333
7	31.738	67.191	49	15.074	16	24	59.272
8	44.952	85.780	49	15.971	16	25	70.136
9	39.667	79.274	49	16.868	16	25	62.235
10	30.857	66.261	49	17.765	16	25	64.210
11	37.024	74.627	49	18.662	16	25	68.160
12	29.095	62.544	49	19.559	16	25	65.198
13	27.333	57.896	49	19.559	17.448	26	55.321
14	31.738	67.191	49	17.765	23.966	26	63.222
15	31.738	67.191	49	18.662	24.690	25	69.148

ExampleSet (334 examples, 0 special attributes, 7 regular attributes)

Gambar 6. Hasil *Preprocessing* Data ISPU 2022

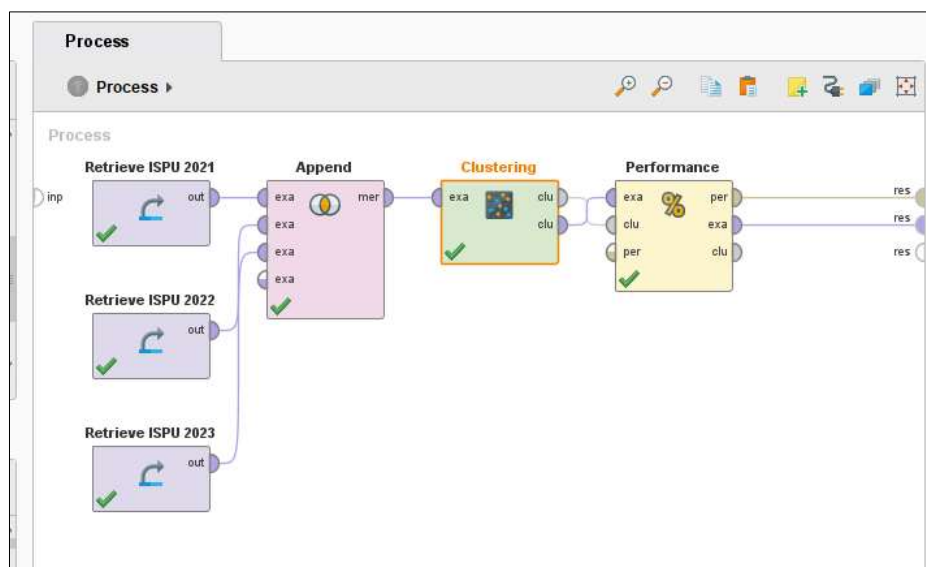
Row No.	PM10	PM2.5	SO2	CO	O3	NO2	HC
1	25	69.053	17	14	7	17.291	1.991
2	20	62.421	17	14	7	17.291	1.991
3	25	69.053	18	14	7	16.468	1.991
4	31	75.684	16	14	7	16.468	1.991
5	30	73.789	23	14	7	17.291	1.991
6	18	58.632	23	14	13	18.114	1.991
7	25	70	18	14	13	17.291	1.991
8	24	68.105	16	14	7	16.468	1.991
9	22.000	66.211	17	14	7	15.646	1.991
10	21	66.211	16	14	6	15.646	1.991
11	38	84.211	15	14	7	15.646	1.991
12	30	75.684	35	14	7	17.291	1.991
13	31	76.632	36	14	9	17.291	1.991
14	22.000	67.158	14	14	9	17.291	1.991
15	20	66.211	16	14	7	16.468	1.991

ExampleSet (357 examples, 0 special attributes, 7 regular attributes)

Gambar 7. Hasil *Preprocessing* Data ISPU 2023

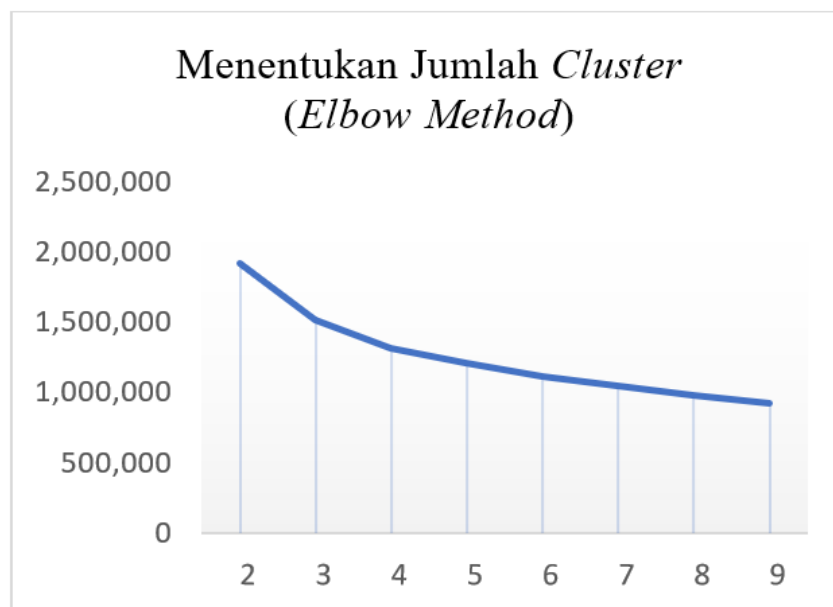
b) Menentukan Jumlah *Cluster*

Menentukan jumlah *cluster* yang optimal untuk *clustering* data Indeks Standar Pencemar udara menggunakan *Elbow Method* dengan membandingkan nilai *average within centroid distance* dari setiap nilai *k*. Berikut tahapan dan hasil penentuannya.



Gambar 8. Menentukan Jumlah Cluster

k	Average Distance
2	2.498.357
3	1.922.800
4	1.519.109
5	1.310.525
6	1.205.209
7	1.116.580
8	1.045.845
9	981.475
10	925.415

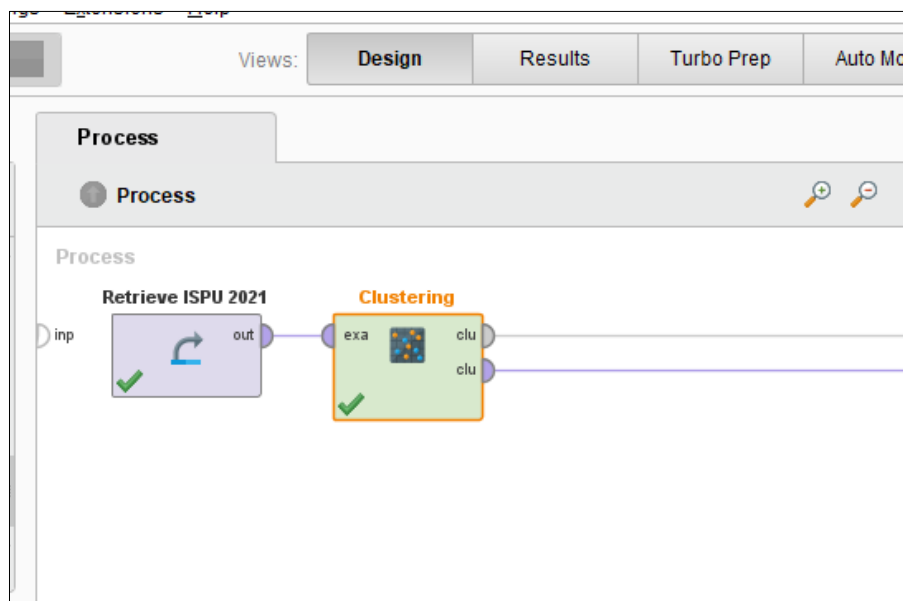


Gambar 9. Elbow Method

Pada *Elbow Method* yang membentuk siku atau cekungan berada di nilai $k=3$. Maka jumlah *cluster* terbaik adalah *cluster* 3 dengan nilai *average distance* 1922.800.

c) Implementasi Clustering K-Means

Berikut langkah-langkah implementasi *clustering* data Indeks Standar Pencemar Udara.

Gambar 10. Implementasi *Clustering* K-Means

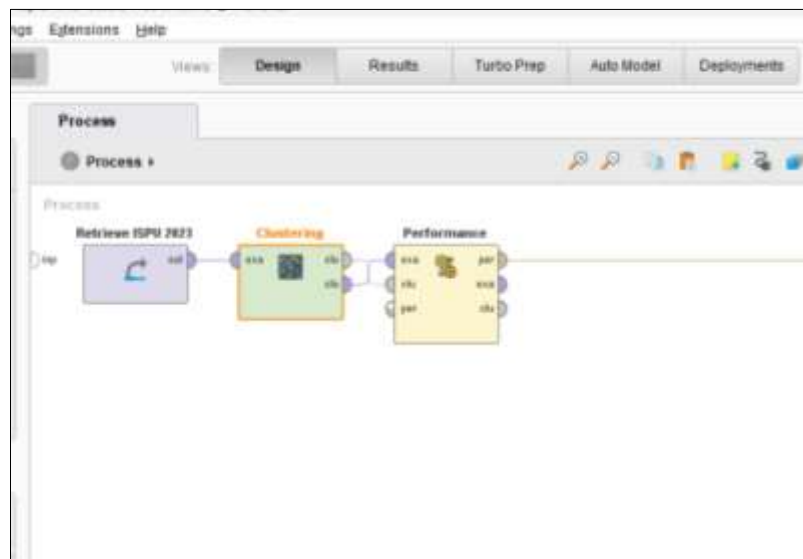
Data yang sudah di *preprocessing* dimasukkan kedalam *canvas* dan tambahkan operator *K-means* dengan nilai k3. Berikut hasil *Cluster Model* nya.

Tabel 5. Cluster Model

Tahun	Jumlah cluster ke-	Jumlah Komponen ISPU	Performance
2021	C0	252	1171.842
	C1	82	1356.069
	C2	7	8265.291
2022	C0	148	1310.432
	C1	171	1204.988
	C2	15	1726.632
2023	C0	15	3613.377
	C1	237	1388.364
	C2	105	2218.466

d) Evaluasi Hasil Clustering

Evaluasi hasil *Clustering* dilakukan dengan cara membanding nilai *Davies Bouldin Index* (DBI)[11]. Berikut hasil dari evaluasi *clustering*.

Gambar 11. Evaluasi Hasil *Clustering*

Evaluasi model terhadap data ISPU dengan memberikan nilai k2 sampai k10 pada K-means untuk melakukan evaluasi cluster K-means dengan DBI dan pada operator performance ubah main criterion menjadi Davies Bouldin. Berikut hasil nya. Tabel 6. Davies Bouldin Indeks Data ISPU.

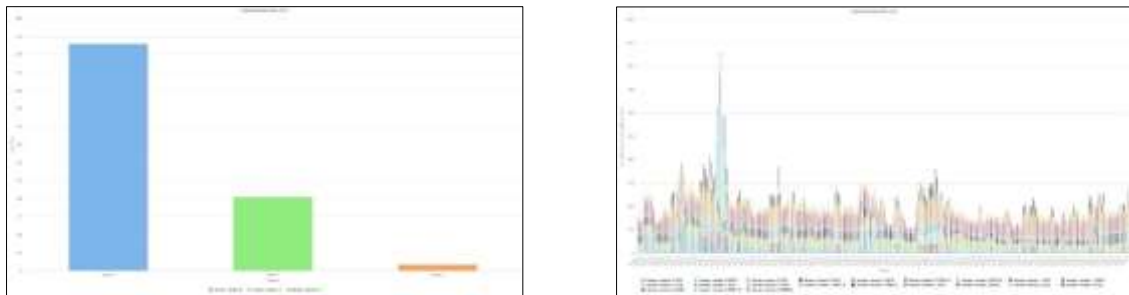
ISPU 2021		ISPU 2022		ISPU 2023	
k	DBI	k	DBI	k	DBI
2	0.399	2	1.478	2	0.744
3	0.881	3	1.036	3	1.011
4	0.744	4	1.182	4	0.968
5	0.892	5	1.235	5	0.987
6	0.896	6	1.288	6	1.042
7	0.890	7	1.189	7	1.012
8	0.991	8	1.177	8	1.048
9	0.958	9	1.093	9	0.990
10	0.962	10	1.062	10	0.984

Melalui perhitungan nilai DBI maka dapat dilihat pada k3 pada masing-masing data ISPU pada tahun 2021, 2022 dan 2023.

e) Visualisasi

Hasil penentuan *cluster* sebelumnya terhadap data ISPU 2021, data ISPU 2022 dan data ISPU 2023.

1) Visualisasi Data ISPU 2021



Gambar 12. Visualisasi data ISPU 2021

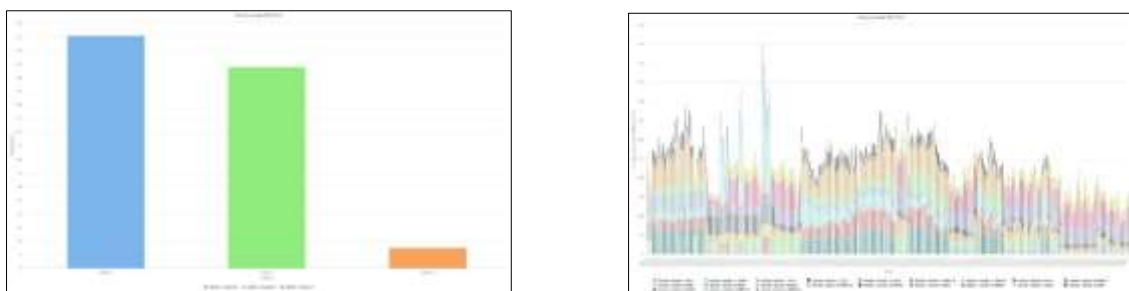
Terdapat tiga cluster yaitu C0 sebanyak 252 data, C1 sebanyak 82 data dan C2 sebanyak 7 data dengan total 341 data dengan dengan konsentrasi tertinggi adalah PM_{2.5} sebesar 429. Maka di dapatlah cluster model sebagai berikut.

Tabel 7. Cluster Model dan Performance data ISPU 2021

No	Cluster ke-	Cluster model	Performance	Kategori Komponen
1	C0	252	1171.842	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
2	C1	82	1356.069	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
3	C2	7	8265.291	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.

Maka dapat dilihat performance C0 merupakan cluster terbaik pada data ISPU 2021 dengan performance 1171.842.

2) Visualisasi Data ISPU 2022



Gambar 13. Visualisasi data ISPU 2022

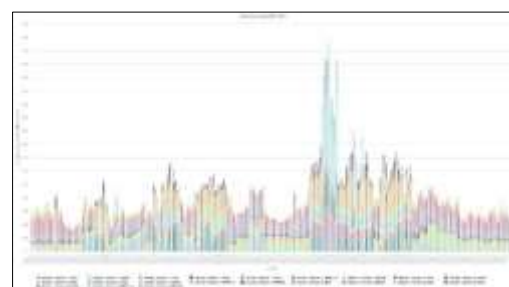
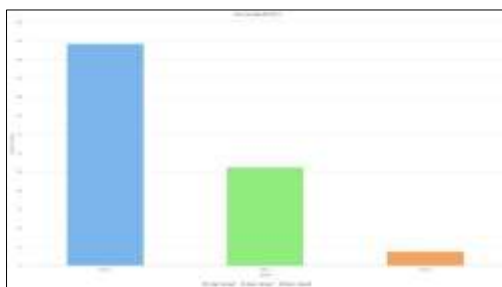
Terdapat 3 cluster yaitu C0 sebanyak 148 data, C1 sebanyak 171 data dan C2 sebanyak 15 data dengan total 334 data dengan konsentrasi tertinggi adalah PM₁₀ sebesar 126. Maka di dapatlah cluster model sebagai berikut.

Tabel 8. Cluster Model dan Performance data ISPU 2022

No	Cluster ke-	Cluster model	Performance	Kategori Komponen
1	C0	148	1310.432	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
2	C1	171	1204.988	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
3	C2	15	1726.632	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.

Maka dapat dilihat performance cluster terbaik adalah C1 dengan performance 1204.988.

3) Visualisasi Data ISPU 2023



Terdapat 3 cluster yaitu C0 sebanyak 15 data, C1 sebanyak 237 data dan C2 sebanyak 105 data dengan konsentrasi tertinggi adalah PM₁₀ sebesar 154. Maka di dapatlah cluster model sebagai berikut.

Tabel 9. Cluster Model dan Performance data ISPU 2023

No	Cluster ke-	Cluster model	Performance	Kategori Komponen
1	C0	15	3613.377	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
2	C1	237	1388.364	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.
3	C2	15	2218.466	PM ₁₀ , PM _{2.5} , SO ₂ , CO, O ₃ , NO ₂ , HC.

Maka dapat dilihat performance cluster terbaik adalah C1 dengan nilai performance 1388.364.

4) Hasil Clustering Data ISPU Tahun

Tabel 10. Hasil Clustering Data ISPU Tahun

Tahun	Jumlah <i>cluster</i> ke-	Jumlah Komponen ISPU	<i>Performance</i>	DBI
2021	C0	252	1171.842	0.881
	C1	82	1356.069	
	C2	7	8265.291	
2022	C0	148	1310.432	1.036
	C1	171	1204.988	
	C2	15	1726.632	
2023	C0	15	3613.377	1.011
	C1	237	1388.364	
	C2	105	2218.466	

Dari tabel hasil *clustering* data ISPU diatas. Maka dihasilkan tiga *cluster* untuk setiap data pertahunnya. Tahun 2021 C0-252, C1-82 dan C2-7, tahun 2022 C0-148, C1-171 dan C2-15, tahun 2023 C0-15, C1-237 dan C2-105..

D. KESIMPULAN

Kesimpulan

Berdasarkan hasil pemodelan, evaluasi dan visualisasi data ISPU di Kota Pontianak peneliti menarik kesimpulan sebagai berikut:

1. Algoritma *K-means* berhasil mengelompokkan data ISPU di Kota Pontianak berdasarkan konsentrasi polutan menjadi 3 *cluster*, seperti polutan PM₁₀, PM_{2.5}, SO₂, CO, O₃, NO₂, HC dengan menggunakan *Elbow Method* dalam menentukan *cluster* terbaik sehingga didapatkan nilai *performance* untuk *k3* yaitu 1922.800.
2. Berdasarkan evaluasi kinerja hasil *clustering* menggunakan *Davies Bouldin Indeks* (DBI) menunjukan nilai DBI untuk *k3* terhadap data ISPU pertahun yaitu pada tahun 2021 nilai DBI sebesar 0.881, tahun 2022 nilai DBI sebesar 1.036 dan tahun 2023 nilai DBI sebesar 1.011.

3. Visualisasi hasil *clustering* menunjukkan konsentrasi polutan tertinggi pertahunnya yang berpotensi menjadi pencermar udara di Kota Pontianak yaitu pada tahun 2021 polutan PM_{2.5} dengan konsentrasi 429, tahun 2022 polutan PM₁₀ dengan konsentrasi 126 dan tahun 2023 polutan PM₁₀ dengan konsentrasi 154.

Saran

Berdasarkan hasil penelitian yang diperoleh dari data di lapangan dan pemodelan *clustering*, peneliti menyampaikan beberapa saran yang diharapkan dapat bermanfaat bagi kemajuan pendidikan secara umum. Adapun saran yang diajukan oleh peneliti adalah sebagai berikut :

1. Memperluas analisis dengan menggunakan algoritma data mining lainnya atau metode *machine learning* yang lebih kompleks untuk melihat apakah hasil yang lebih akurat dapat dicapai.
2. Penerapan kebijakan pengendalian emisi yang lebih ketat, khususnya pada sektor-sektor yang berkontribusi besar terhadap polutan PM₁₀ dan PM_{2.5} seperti industri dan transportasi.

DAFTAR PUSTAKA

- T. Rizqiyah and I. Rosyida, 'Analisis Cluster Tingkat Kualitas Udara Ambien Jalan Raya di Jawa Tengah Tahun', *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 4, pp. 560–564, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- PERMEN P.14/MENLHK/SETJEN/KUM.1/7/2020, 'Indek Standar Pencemar Udara', in *PERATURAN MENTERI LINGKUNGAN HIDUP DAN KEHUTANAN*, 2020. Accessed: Jul. 08, 2024. [Online]. Available: https://ditppu.menlhk.go.id/portal/uploads/news/1600940556_P_14_2020_ISPU_menlhk_07302020074834.pdf
- S. Khairunnisa and M. I. Jambak, 'Pengelompokan Cuaca Kota Palembang Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Pola Karakteristik Cuaca', *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, p. 2352, Oct. 2022, doi: 10.30865/mib.v6i4.4810.

- A. Rifqi and R. T. Aldisa, 'Penerapan Data Mining Untuk Clustering Kualitas Udara', *Jurnal Sistem Komputer dan Informatika (JSON) Hal: 289–*, vol. 297, no. 2, 2023, doi: 10.30865/json.v5i2.7145.
- F. N. Dhewayani *et al.*, 'Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM', *Jurnal Teknologi dan Informasi*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.
- K. Hutan, D. Lahan, and P. Y. Utami, 'ANALISIS CLUSTERING K-MEANS PADA PENGELOMPOKKAN TITIK PANAS', *Jurnal Pendidikan Informatika dan Sains*, vol. 12, no. 1, pp. 165–172, 2023, doi: 10.31571/saintek.v12i1.6001.
- A. Firmansyah *et al.*, 'Pengelompokan Titik Wilayah di Provinsi Daerah Istimewa Yogyakarta Berdasarkan Kualitas Udara Menggunakan Algoritma Fuzzy C-Means', 2020.
- M. Rosadi, D. A. Nurhasanah, and M. S. Hasibuan, 'Clustering Panjang Ruas Jalan di BBPJN Sumut Menggunakan Algoritma K-Means Clustering The Length Of Road Sections In BBPJN Sumut Using The K-Means Algorithm', *JOURNAL OF COMPUTER SCIENCE AND INFORMATICS ENGINEERING (CoSIE)*, vol. 02, no. 1, p. 2023, 2023, [Online]. Available: <http://creativecommons.org/licenses/by-sa/4.0/>
- M. Sholeh and D. Andayati, 'Penerapan Metode Clustering dengan Algoritma K-Means Pada Pengelompokan Indeks Prestasi Akademik Mahasiswa', *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 6, no. 1, pp. 51–60, 2023.
- I. Hidayat, E. Darnila, and Y. Afrillia, 'Clustering Zonasi Daerah Rawan Bencana Alam di Kabupaten Mandailing Natal menggunakan Algoritma K-Means', *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 3, pp. 1218–1226, Jul. 2023, doi: 10.33379/gtech.v7i3.2880.
- D. Rohman, R. Annisa, D. I. Efendi, and D. Solahudin, 'CLUSTERING BENCANA ALAM MENGGUNAKAN K-MEANS PADA WILAYAH JAWA BARAT', 2024. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK55>.