

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini, akan dibahas hasil analisis yang telah dilakukan terkait dengan sumber emisi udara dan tracing material input dalam industri powerplant serta oil & gas distribution pada perusahaan ailesh menggunakan metode k-means clustering. Proses analisis ini dilakukan untuk analisis data emisi dan material yang dikumpulkan oleh Ailesh dari berbagai unit operasionalnya. Data ini akan dianalisis menggunakan metode *K-Means Clustering*. Data yang digunakan dalam penelitian ini diperoleh dengan mengajukan kerjasama dengan pihak Perusahaan ailesh yaitu menyajikan data hasil analisis sesuai kebutuhan Perusahaan dalam bentuk dashboard. Data tersebut berupa data dari sektor Oil & Gas Distribution dan data Power Plant, yang mencakup data material input serta data output. Komunikasi dan pengumpulan data dilakukan melalui platform daring seperti whatsapp, google meet, dan aplikasi zoom meeting.

Metode K-Means clustering digunakan untuk mengelompokkan atau mengidentifikasi kelompok sumber emisi dan material berdasarkan karakteristik tertentu. Hasil yang didapatkan dari analisis ini diharapkan dapat memberikan manfaat bagi Perusahaan ailesh untuk menangani atau mengurangi material pada industri tersebut yang menjadi sumber emisi udara sehingga berdampak terhadap pencemaran lingkungan, dan penelitian ini juga diharapkan dapat bermanfaat bagi institusi, penulis serta masyarakat sekitar.

4.1 Data Collection (Pengumpulan Data)

4.1.1 Pengumpulan Data

Langkah pertama dalam proses data mining adalah pengumpulan data. Data yang digunakan dalam penelitian ini merupakan data dari industri *Power Plant* serta *Oil & Gas Distribution* diperoleh melalui Perusahaan ailesh yang diberikan dalam bentuk excel melalui link google drive dari Perusahaan tersebut. Data yang

diberikan perusahaan berupa data mentah yang kemudian dilakukan tahapan normalisasi ke dalam 1NF. Setelah data dinormalisasi ke dalam 1NF maka data disimpan dalam bentuk file csv, data yang dikumpulkan dari sektor power plant dan oil & gas distribution sejumlah 1499 data dan 13 kolom (Atribut) yang nantinya akan dianalisis menggunakan teknik data mining.

4.2 Data Understanding (Pemahaman Data)

Setelah data terkumpul, langkah berikutnya adalah untuk memahami data tersebut. Pada tahap ini, dilakukan eksplorasi awal terhadap data untuk memahami struktur, karakteristik, dan kualitasnya. Tujuannya adalah untuk mengetahui informasi apa yang ada dalam data, serta untuk mendeteksi masalah seperti tipe data yang tidak sesuai, *missing values*, *duplikasi*, atau *inkonsistensi*.

4.2.1 Menyiapkan Data dan Library

Bila dataset yang dibutuhkan sudah tersedia, tahap selanjutnya adalah menyiapkan library pada google colab untuk memudahkan dalam pengolahan data.

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.cluster import KMeans
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
import seaborn as sns
```

Gambar 4.1 *Code Import Library*

4.2.2 Mengimpor Data Ke Google Colab

Selanjutnya Pada tahapan ini penulis akan mengimpor data penelitian yang telah disimpan dalam bentuk csv tersebut di google colab. Kode di di bawah ini berfungsi untuk mengimpor dataset penelitian ke google colab.

```
# Menyambungkan Google Colab ke Google Drive agar dapat mengakses Dataset / DataFrame  
  
from google.colab import drive # library untuk mengakses Google Drive dari Google Collab  
drive.mount('/content/drive') # Menghubungkan ke google drive  
file_path = '/content/drive/My Drive/Data Penelitian/Analisis Doom spending.csv' # Alamat File data  
df = pd.read_csv(file_path) #Membaca File
```

Gambar 4. 2 *Code Import Data*

Hal lain yang dapat dilakukan pada dataset adalah mendapatkan informasi terkait jumlah kolom, jumlah seluruh data yang terdapat pada dataset.

```
df.shape  
  
(1499, 14)
```

Gambar 4. 3 Jumlah Baris Kolom

Dapat dilihat pada gambar 4.3 bahwa data penelitian terdapat 1499 baris dan 14 kolom, dengan sebagian besar kolom berisi data kategorikal atau teks (dengan tipe data object), kecuali kolom value yang berisi data numerik (dengan tipe data Float).

4.2.3 Mengoreksi Kesalahan Tipe data

Penyesuaian tipe data adalah langkah penting dalam analisis data, hal ini untuk memastikan data siap diproses. Dengan tipe data yang tepat, analisis yang dilakukan akan lebih akurat.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1499 entries, 0 to 1498
Data columns (total 14 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   industry                             1499 non-null   object
1   sector                               1499 non-null   object
2   location                             1499 non-null   object
3   unit_process                         1499 non-null   object
4   input_output                         1499 non-null   object
5   category                             1499 non-null   object
6   subcategory                          1499 non-null   object
7   material_subcategory                 1499 non-null   object
8   material_rename                     1499 non-null   object
9   material                             1499 non-null   object
10  value                               1499 non-null   object
11  unit                                1499 non-null   object
12  persentase                           170 non-null    object
13  unit_process_subcategory             1499 non-null   object
dtypes: object(14)
memory usage: 164.1+ KB
```

Gambar 4. 4 Penjelasan Dataset

Pada gambar 4.4 menunjukkan bahwa pada kolom value tipe data nya masih dalam bentuk kategorikal (object), seharusnya tipe data yang sesuai yaitu tipe data numerik(float) karena pada kolom value terdapat nilai atau jumlah bahan yang digunakan dari masing-masing industry dan sektor. Tipe data float digunakan untuk merepresentasikan bilangan desimal atau pecahan dalam pemrograman oleh karena itu mengapa tipe data float cocok untuk kolom value karena berisi nilai desimal atau pecahan. Berikut kode untuk merubah tipe data pada kolom *value* dari tipe data object menjadi tipe data float:

```
# remove value
df = df[df.value != '#REF!']

# convert dtype
df['value'] = df['value'].astype(float)
```

Gambar 4. 5 Code Mengubah Tipe Data

Kode tersebut merupakan kode merubah tipe data pada kolom value dari bentuk kategorikal (object) menjadi numerik (float), setelah menjalankan kode tersebut dapat dilihat pada gambar berikut tipe data dari masing-masing kolom sudah sesuai.

```
<class 'pandas.core.frame.DataFrame'>
Index: 1498 entries, 0 to 1498
Data columns (total 14 columns):
#   Column                Non-Null Count  Dtype
---  -
0   industry              1498 non-null  object
1   sector                1498 non-null  object
2   location              1498 non-null  object
3   unit_process          1498 non-null  object
4   input_output          1498 non-null  object
5   category              1498 non-null  object
6   subcategory           1498 non-null  object
7   material_subcategory  1498 non-null  object
8   material_rename       1498 non-null  object
9   material              1498 non-null  object
10  value                 1498 non-null  float64
11  unit                  1498 non-null  object
12  persentase            170 non-null   object
13  unit_process_subcategory 1498 non-null  object
dtypes: float64(1), object(13)
memory usage: 175.5+ KB
```

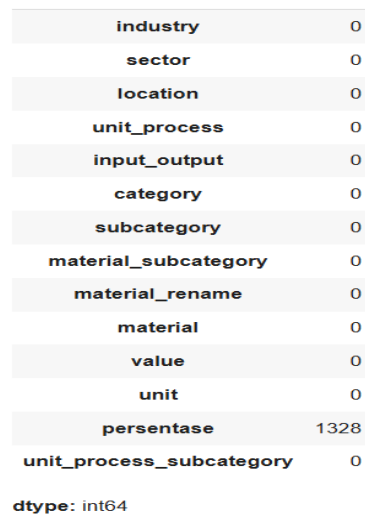
Gambar 4. 6 Penjelasan Tipe Data

4.3 Data *Cleaning* (Pembersihan Data)

Tujuan pembersihan data adalah untuk menghapus data yang dianggap tidak seimbang, yang tidak lengkap atau data yang tidak sesuai.

4.3.1 Menangani *Missing Values*

Pada tahap ini dilakukan pemeriksaan pada dataset yang bertujuan mengetahui nilai yang rumpang atau missing values untuk menghindari overfitting pada model, data yang rumpang sangat sensitif terhadap performa model. Oleh karena, itu harus dilakukan pemeriksaan missing value.



industry	0
sector	0
location	0
unit_process	0
input_output	0
category	0
subcategory	0
material_subcategory	0
material_rename	0
material	0
value	0
unit	0
persentase	1328
unit_process_subcategory	0

dtype: int64

Gambar 4. 7 *Chek Missing Value*

Setelah dilakukan pemeriksaan missing value pada gambar 4.7 menunjukkan informasi dalam dataset yaitu hanya terdapat satu kolom yang memiliki missing value (Tidak ada nilainya/null), yaitu pada kolom persentase terdapat 1328 missing value.

```
# Drop the 'persentase' column due to many missing values
df = df.drop('persentase', axis=1)

# Verify the column has been removed
df.columns
```

Gambar 4.8 *Code Drop Column*

Dapat dilihat pada gambar 4.8 kode drop kolom yang terdapat missing value, langkah selanjutnya yaitu menghapus kolom persentase karena terdapat banyak missing value yaitu 1328. Hal ini merupakan langkah yang penting karena kolom persentase tidak digunakan saat analisis tetapi jika kolom persentase tidak dihapus akan sangat berpengaruh pada data yang akan di analisis.

Setelah menghapus kolom persentase yang terdapat 1328 *missing value* maka dapat dilihat bahwa missing value sudah selesai di tangani yaitu dengan menghapus kolom persentase yang tidak digunakan dalam analisis sehingga jumlah kolom pada data berkurang menjadi 13 kolom dan jumlah data tetap 1499 data. Pada kolom yang lain dapat dilihat bahwa tidak terdapat *missing value* yaitu ditandai dengan nilai 0.

	0
industry	0
sector	0
location	0
unit_process	0
input_output	0
category	0
subcategory	0
material_subcategory	0
material_rename	0
material	0
value	0
unit	0
unit_process_subcategory	0

dtype: int64

Gambar 4.9 *Setelah Penanganan Missing Value*

4.3.2 Memeriksa dan Menghapus Kolom Duplikat

Dalam penelitian ini, salah satu langkah penting yang dilakukan adalah menangani data duplikat yang ditemukan dalam dataset. Berikut kode python untuk menampilkan data yang duplikat :

```
# prompt: cek data duplikat yang mempunyai nilai sama setiap baris dengan kolom yang lain

# check for rows where all columns have the same value
duplicate_rows = df[df.apply(lambda row: row.nunique() == 1, axis=1)]

print("Rows with duplicate values across all columns:")
duplicate_rows
```

Gambar 4. 10 *Code Chek Data Duplikat*

Setelah dilakukannya pemeriksaan data duplikat pada data yang akan dianalisis, hasil pengemeriiksaan tidak terdapat data yang duplikat, sehingga penulis tidak melakukan penanganan dan penghapusan data duplikat.

4.4 Data Eksplorasi (Data Exploration)

Dalam case ini, sangat penting untuk memahami industri-industri yang memiliki kontribusi signifikan terhadap emisi karbon global, yaitu industri Power Plant dan Oil & Gas Distribution. Industri oil & gas distribution diperkirakan menyumbang besar dalam emisi karbon global. Hal ini karena proses ekstraksi, pengolahan, dan penggunaan produk minyak dan gas akan menyebabkan pelepasan karbon dioksida ke atmosfer.

Dalam dampak yang panjang, pelepasan emisi karbon dapat mempengaruhi perubahan iklim global, perubahan siklus cuaca dan suhu, seperti yang telah banyak kita rasakan. Namun, industri oil dan gas masih menjadi sumber energi global yang banyak digunakan di berbagai negara meskipun saat ini telah banyak digalakkan transisi ke sumber energi terbarukan. Begitu juga dengan industri power plant atau pembangkit listrik.

Power plant juga menjadi salah satu sumber emisi gas rumah kaca terbesar. Hal ini dikarenakan, berbagai jenis pembangkit masih banyak menggunakan batu bara, gas alam, dan minyak, sehingga banyak menghasilkan emisi dalam aktivitasnya. Melihat urgensi industri oil & gas serta power plant yang penting ini, perlu dilakukan analisis pada kedua industri tersebut agar dapat diketahui solusi yang paling tepat untuk membantu mengurangi dampak emisi dan melaksanakan transformasi hijau yang berkelanjutan, sehingga proses industri dan bisnis yang dijalankan memiliki aspek lingkungan, dekarbonisasi, sirkularitas, dan transformasi ke berkelanjutan.

4.4.1 Analisis Sumber Emisi Udara

Tahapan analisis sumber emisi udara yaitu mengidentifikasi sumber emisi yang dihasilkan dari lokasi- lokasi pada data yang diberikan perusahaan ailesh. Emisi yang difokuskan yaitu emisi udara dari material seperti carbon dioxid (CO₂), sulfur dioxid (SO₂) dan metan yang menyebabkan pemanasan global, perubahan iklim dan semua yang membahayakan lingkungan. Sektor industry yang difokuskan yaitu industry powerplant serta oil & gas distribution, Dimana keadua industry ini diketahui menyumbang besar dalam emisi carbon. Hal ini karena proses ekstraksinya, pengolahan dan penggunaan produk oil dan gas serta berbagai jenis input material yang digunakannya menghasilkan emisi yang tinggi.

1. Emisi udara yang dihasilkan masing-masing industry

	industry	value
0	Oil & Gas Distribution	4.036136e+09
1	Power Plant	6.402562e+07

Gambar 4. 11 Rata-rata Jumlah Emisi

Berdasarkan penghitungan nilai *emission to air* (Emisi ke udara) yang telah dilakukan pada industry oil & gas distribution serta power plant, dapat dilihat pada gambar 4.15 ditemukan emisi sebanyak :

- Oil & Gas Distribution: 4,036,136,000 ton.
- Power Plant: 64,025,620 ton.

Dapat disimpulkan bahwa kegiatan pada industri Oil & Gas Distribution menghasilkan emisi ke udara lebih banyak daripada industri Power Plant.

c. *Insight Oil & Gas Distribution* (pemahaman yang diperoleh dari industri minyak dan gas distribusi)

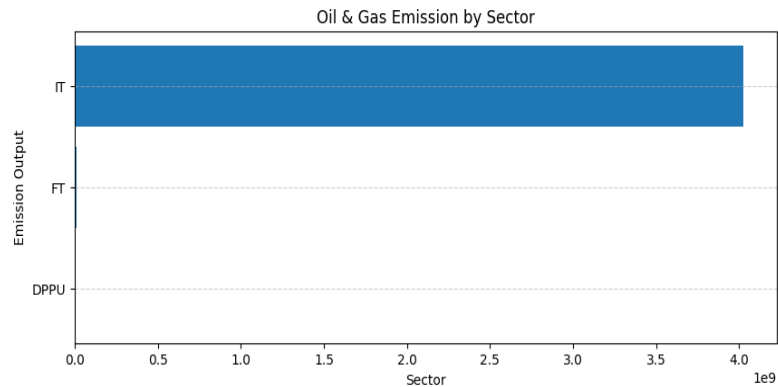
Pemahaman analisis industri oil & gas distribusi untuk melihat informasi mengenai data emisi udara secara spesifik, pada tahapan ini kode yang di gunakan berfungsi untuk menganalisis dan merangkum data emisi udara secara spesifik untuk *industry oil & gas distribution* saja. Kemudian menghitung total emisi untuk setiap sub-sektor di dalamnya, mengurutkan hasilnya, dan mempersiapkan data tersebut untuk visualisasi lebih lanjut.

	sector	value
0	IT	4.025105e+09
1	FT	1.102167e+07
2	DPPU	9.049241e+03

Gambar 4. 12 Jumlah Emisi Berdasarkan Sektor

Berdasarkan perhitungan maka didapat hasil dari perhitungan yaitu total emisi berdasarkan sektor pada *industry oil & gas distribution* yang paling banyak yaitu pada sektor IT (Injection Terminal) sebanyak 4.025.105.000 (4 Milyar) ton, sedangkan sektor FT (Fuel Terminal) sebanyak 11.021.670 (11 Juta) ton dan sektor DPPU (Depot Pengisian Pesawat Udara) sebanyak

9.049.24 (9 juta) ton. IT (Injection Terminal) ialah dimana minyak atau gas disuntikkan ke dalam pipa atau jaringan untuk didistribusikan ke tujuan akhir, IT sering digunakan untuk memindahkan produk dari fasilitas penyimpanan atau produksi ke sistem pipa utama.



Gambar 4. 13 Visualisasi Emisi Berdasarkan *Sector*

Setelah mengetahui sektor yang menghasilkan emisi paling banyak akan dilakukan analisis lebih dalam lagi untuk melihat unit proses yang menghasilkan emisi terbesar .

	unit_process	value
0	Use Stage	4.036124e+09
1	Storage Tank	5.695127e+03
2	Pompa	2.089281e+03
3	Mobil Tangki	1.782266e+03
4	Gas Turbine	1.180000e+03

Gambar 4. 14 Jumlah Emisi Berdasarkan Unit Proses

Berdasarkan analisis jumlah emisi berdasarkan unit proses, gambar 4.18 menunjukkan bahwa terdapat 5 unit proses penghasil emisi udara tertinggi yaitu:

- Use Stage: 4,036,124,000 ton

Use Stage adalah proses ketika bahan bakar digunakan (misalnya dibakar di mesin kendaraan atau generator), di sinilah emisi gas rumah kaca (seperti CO₂) dan polutan lain (seperti NO_x) dihasilkan paling banyak.

- Storage Tank: 5,695.127 ton

Storage Tank (Tangki Penyimpanan) menyimpan bahan bakar cair seperti minyak dan gas. Saat bahan bakar ini disimpan, uap yang mudah menguap (VOC) bisa keluar ke udara, ini adalah salah satu sumber emisi.

- Pompa: 2,089.281 ton

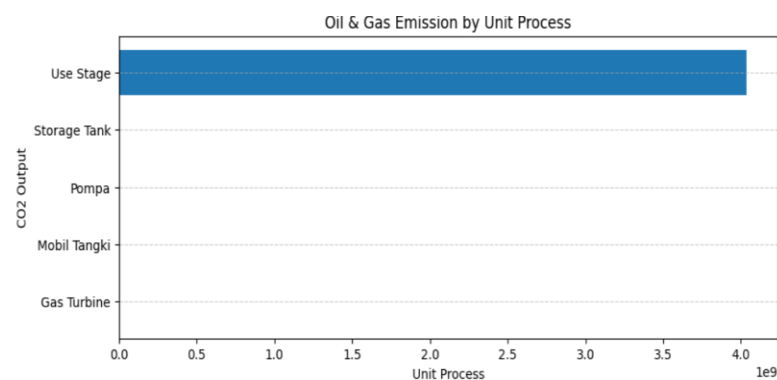
Pompa adalah Alat yang memindahkan cairan atau gas dari satu tempat ke tempat lain, seperti memindahkan bahan bakar dari tangki ke mesin.

- Mobil Tangki: 1,782.266 ton

Mobil Tangki merupakan Truk khusus yang digunakan untuk mengangkut bahan bakar atau cairan kimia dalam jumlah besar. Kadang bisa menyebabkan uap bahan bakar keluar saat pengisian atau pengosongan.

- Gas Turbine: 1,180.000 ton

Mesin yang mengubah energi dari gas panas hasil pembakaran menjadi energi listrik. Gas panas menggerakkan turbin, dan turbin menggerakkan generator listrik.



Gambar 4.15 Visualisasi Emisi Berdasarkan Unit Proses

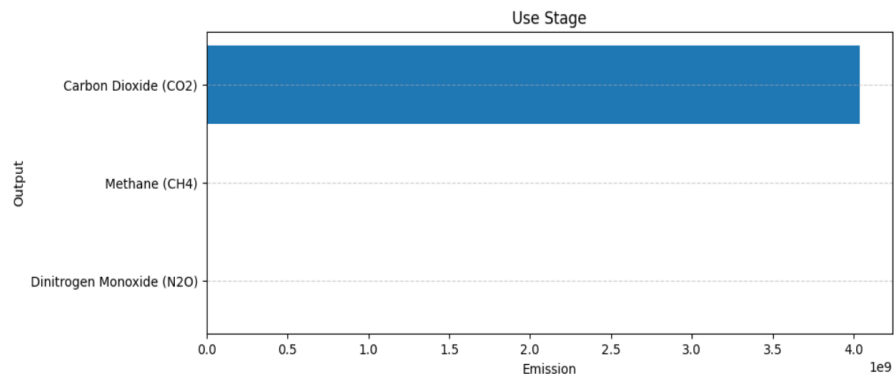
Proses Use Stage menjadi unit proses yang menghasilkan emisi paling banyak di antara unit proses lainnya dalam industri *Oil & Gas Distribution*, proses Use Stage merujuk pada tahap di mana produk minyak dan gas digunakan oleh konsumen akhir. Proses ini melibatkan berbagai aktivitas dan dapat memiliki dampak signifikan terhadap lingkungan, terutama dalam hal emisi gas rumah kaca dan polutan lainnya. Selanjutnya dilakukan analisis material emisi yang dihasilkan oleh proses Use Stage. Berikut dapat dilihat pada gambar dibawah ini :

	material_rename	value
0	Carbon Dioxide (CO2)	4.035797e+09
1	Methane (CH4)	2.852122e+05
2	Dinitrogen Monoxide (N2O)	4.170401e+04

Gambar 4. 16 Emisi dari Proses Use Stage

Dari unit proses Use Stage, diketahui ada beberapa jenis material emisi yang dilepaskan, di antaranya:

- Carbon Dioxide (CO2): 4,035,797,000 ton
- Methana (CH4): 285,212 ton
- Dinitrogen Monoxide (N2O): 41,704 ton



Gambar 4. 17 Visualisasi Emisi Proses Use Stage

Carbon Dioxide atau CO₂ menjadi material emisi yang paling banyak dilepaskan dari proses Use Stage ini. Proses use stage dalam industri oil & gas distribution melepaskan banyak emisi CO₂ karena beberapa alasan yaitu Selama proses pembakaran use stage, produk oil & gas seperti bensin, diesel, dan gas alam dibakar untuk menghasilkan energi dalam jumlah yang besar. Proses pembakaran ini secara kimiawi mengubah karbon dalam bahan bakar menjadi karbon dioksida (CO₂) sangat besar, yang kemudian dilepaskan ke atmosfer. Sedangkan, di sisi lain alternatif yang ada hanya terbatas, meskipun ada peningkatan penggunaan energi terbarukan, seperti angin dan matahari, tetapi sektor industri dan transportasi masih banyak bergantung pada bahan bakar fosil.

- d. *Insight Power plant* (pemahaman yang diperoleh dari industri pembangkit listrik)

	sector	value
0	PLTG PLTGU	6.059012e+07
1	PLTU	3.435495e+06
2	PLTA	6.736023e-01

Gambar 4. 18 Jumlah Emisi Berdasarkan Sektor

Pada gambar 4.22 dapat dilihat bahwa sektor PLTG/PLTGU (Pembangkit listrik tenaga gas / Pembangkit listrik tenaga gas uap) menghasilkan emisi paling banyak daripada PLTU (Pembangkit listrik tenaga uap) dan PLTA (Pembangkit listrik tenaga air), sektor PLTG/PLTGU dengan total emisi yang diperoleh yaitu sebanyak 60.590.120 ton. Hal ini dikarenakan sektor PLTG dan PLTGU mengandalkan proses pembakaran bahan bakar fosil untuk menghasilkan listrik. Proses pembakaran pada sektor PLTG dan PLTGU menghasilkan emisi CO₂

dan polutan lainnya seperti nitrogen oksida (NO_x), sulfur dioksida (SO₂), dan partikel-partikel halus (PM).

Setelah mengetahui sektor mana yang menghasilkan emisi paling banyak pada *industry power plant* maka akan dilakukan analisis lebih lanjut untuk melihat unit proses yang menghasilkan emisi terbesar.

	unit_process	value
0	Hrsg	6.041298e+07
1	Boiler	3.435494e+06
2	Gas Turbine	1.748353e+05
3	Gas Turbine Generator	2.302392e+03
4	Water Treatment Plant	1.349040e+00

Gambar 4. 19 Tingkat Emisi Berdasarkan Unit Proses

Berdasarkan informasi pada gambar 4.23, diketahui terdapat beberapa unit proses dalam industri *Power Plant*, yaitu:

- Hrsg: 60,412,980 ton

HRSG merupakan alat yang menangkap panas dari gas buang turbin gas untuk menghasilkan uap. Uap ini kemudian bisa digunakan untuk pembangkit listrik tambahan, jadi energi panas tidak terbuang sia-sia.

- Boiler: 3,435,494 ton

Boiler adalah alat yang digunakan untuk memanaskan air hingga berubah menjadi uap. Uap ini kemudian digunakan untuk menggerakkan mesin atau turbin agar bisa menghasilkan listrik atau menjalankan proses industri. Boiler biasanya membakar bahan bakar seperti gas atau batu bara untuk menghasilkan panas.

- Gas Turbine: 174,835.3 ton

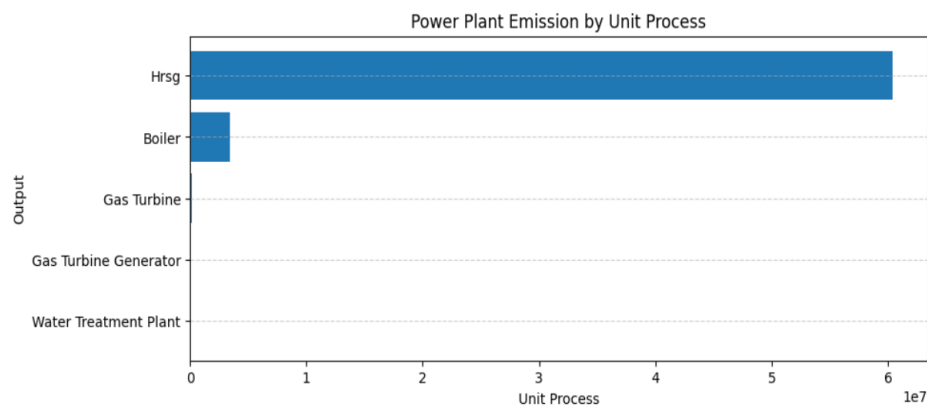
Gas turbine adalah mesin pembakaran internal yang mengubah energi dari bahan bakar (biasanya gas atau bahan bakar cair) menjadi energi mekanik melalui proses pembakaran bertekanan tinggi. Energi mekanik ini digunakan untuk menggerakkan generator listrik, kompresor, atau sistem propulsi seperti pada pesawat terbang.

- Gas Turbine Generator: 2,302.392 ton

Gas turbine generator adalah mesin yang mengubah energi dari gas panas hasil pembakaran menjadi energi listrik. Gas panas menggerakkan turbin, dan turbin menggerakkan generator listrik.

- Water Treatment Plant: 1.34904 ton

Water Treatment Plant adalah fasilitas yang mengolah air limbah agar aman dibuang ke lingkungan. Proses ini membersihkan air dari bahan berbahaya dan polutan.



Gambar 4. 20 Visualisasi Emisi Berdasarkan Unit Proses

HRSG menjadi unit proses yang banyak menghasilkan emisi ke udara dalam aktivitasnya. HRSG (Heat Recovery Steam Generator) merupakan jenis penukar panas pemulihan energi yang memanfaatkan panas dari aliran gas

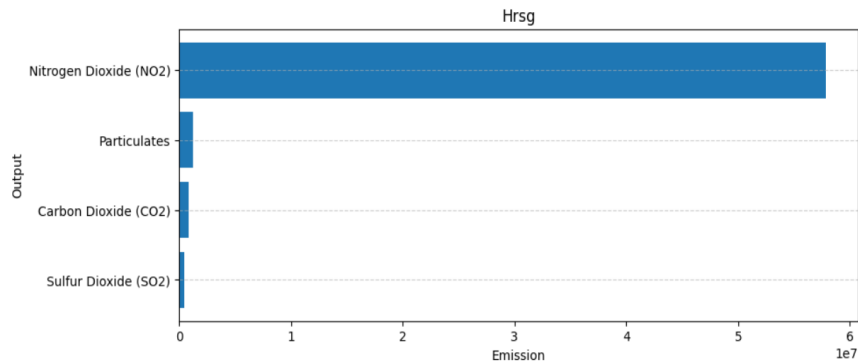
panas, biasanya dari knalpot turbin gas, untuk menghasilkan uap yang dapat digunakan dalam turbin uap atau untuk proses industri lainnya.

	material_rename	value
0	Nitrogen Dioxide (NO ₂)	5.785447e+07
1	Particulates	1.225807e+06
2	Carbon Dioxide (CO ₂)	8.987772e+05
3	Sulfur Dioxide (SO ₂)	4.339268e+05

Gambar 4. 21 Material Emisi dari Proses HRSG

Dari proses HRSG, diketahui ada beberapa jenis material emisi yang dilepaskan, di antaranya:

- Nitrogen Dioxide (NO₂): 57,854,470 ton
Nitrogen Dioksida (NO₂) adalah salah satu komponen utama emisi udara dari kendaraan dan industri. Gas ini mencemari udara dan berkontribusi pada pembentukan hujan asam dan ozon permukaan yang berbahaya.
- Particulates: 1,225,807 ton
Partikulat (PM) adalah partikel halus dalam emisi udara yang berasal dari pembakaran, debu, dan proses industri. Partikulat berbahaya karena bisa terhirup dan menimbulkan gangguan kesehatan serius.
- Carbon Dioxide (CO₂): 898,777.2 ton
Karbon Dioksida (CO₂) merupakan emisi gas rumah kaca utama dari pembakaran bahan bakar fosil. Meskipun tidak beracun langsung, CO₂ memicu pemanasan global dan memperburuk kualitas udara secara tidak langsung.
- Sulfur Dioxide (SO₂): 433,926.8 ton
Sulfur Dioksida (SO₂) adalah komponen emisi udara dari pembakaran bahan bakar yang mengandung sulfur. Gas ini memperburuk pencemaran udara dan menyebabkan hujan asam yang merusak lingkungan.



Nitrogen Dioxide atau NO2 menjadi jenis material paling banyak yang dilepaskan pada proses HRSG.

2. *Tracing Material input (Pelacakan Bahan yang digunakan)*

Analisis tracing material input adalah analisis untuk melacak material apa saja yang menjadi kontributor dalam proses yang menghasilkan emisi udara pada industri. Dengan melakukan pelacakan dapat mengetahui material yang menjadi kontributor terbesar dalam menghasilkan emisi udara, sehingga membantu perusahaan dalam memahami hubungan antara penggunaan material input dengan tingkat emisi udara yang dihasilkan.

a. Proses *Use Stage*

```
# check locations with use stage
df_og_usestage['location'].unique()

array(['Adi Sucipto', 'Boyolali', 'Ahmad Yani', 'Adi Sumarno', 'Makassar',
       'Cilacap', 'Rewulu', 'Bitung'], dtype=object)
```

Gambar 4. 22 Kode cek lokasi proses Use Stage

Dapat dilihat pada gambar 4.26 terdapat delapan tempat yang melakukan proses use stage. Untuk mengetahui kenapa unit process ini menghasilkan emisi, selanjutnya penulis akan melakukan *tracing material input* pada lokasi-lokasi yang melakukan proses use stage.

	subcategory	value
0	Energy	413193.934510
1	Material	268346.664503
2	Diesel	1440.785934
3	Chemical	10.497000

Gambar 4. 23 Emisi Material Proses Use Stage

Pada gambar 4.27 dapat dilihat bahwa diketahui ada beberapa kategori material yang digunakan sebagai input dari lokasi-lokasi yang melakukan proses use stage, di antaranya:

- Material: 206,690,500 kl
Material mengacu pada bahan mentah atau setengah jadi yang digunakan dalam proses produksi, seperti logam, plastik, kayu, dan lainnya. Pengelolaan material yang tidak efisien dapat menyebabkan limbah dan emisi tambahan dalam proses industri.
- Energy: 1,850,105 kWh
Energi dibutuhkan untuk menjalankan mesin, alat produksi, penerangan, dan sistem lainnya. Sumber energi seperti listrik dan bahan bakar fosil menyumbang besar terhadap emisi gas rumah kaca, terutama jika berasal dari pembangkit berbasis batu bara atau minyak.

- Diesel: 1,440.786 t

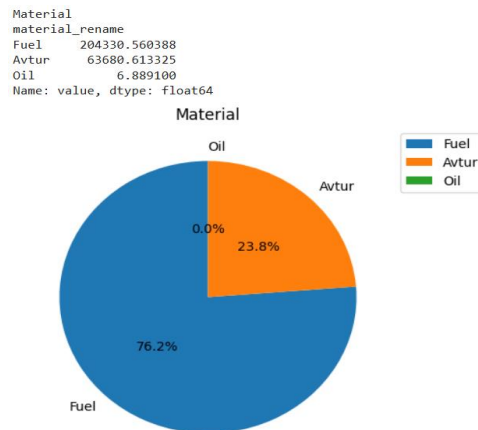
Diesel adalah bahan bakar berbasis minyak yang umum digunakan pada kendaraan berat dan mesin industri. Pembakaran diesel menghasilkan emisi berbahaya seperti NO₂, partikulat, dan CO₂ yang berdampak negatif terhadap kualitas udara dan kesehatan manusia.

- Chemical: 10.5

Chemical (bahan kimia) digunakan luas di industri untuk proses produksi, pengawetan, dan pembersihan. Penggunaan dan pembuangan bahan kimia yang tidak terkendali dapat mencemari udara, air, dan tanah serta membahayakan kesehatan manusia dan lingkungan

Kategori-kategori *input* tersebut dapat di analisis lebih lanjut lagi untuk melihat material apa saja yang digunakan, di antaranya:

1. Material



Gambar 4. 24 Kategori Material Proses Use Stage

- Fuel: 187,340,600 kl

Fuel (bahan bakar) adalah sumber energi utama yang digunakan untuk pembakaran dalam mesin, kendaraan, dan industri. Jenis bahan bakar bisa berupa bensin, solar, gas alam, dan lainnya. Penggunaan bahan

bakar ini menghasilkan emisi gas rumah kaca seperti CO₂ dan polutan lain yang mencemari udara.

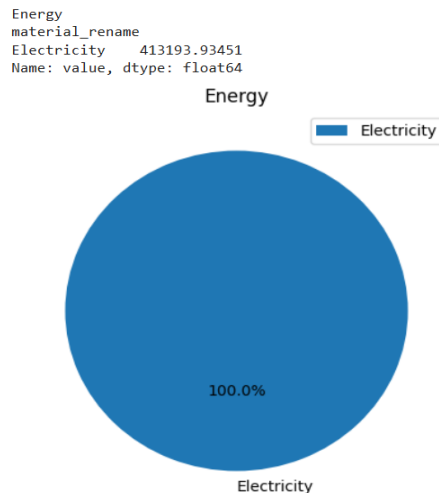
- Avtur: 19,349,550 kl

Avtur (Aviation Turbine Fuel) adalah bahan bakar khusus untuk pesawat terbang jet. Pembakarannya menghasilkan emisi karbon dioksida, nitrogen oksida, dan uap air di ketinggian, yang berkontribusi pada pemanasan global dan gangguan atmosfer.

- Oil: 6.8891 kl

Oil (minyak) dapat merujuk pada minyak bumi atau pelumas. Dalam konteks energi, minyak digunakan sebagai bahan bakar atau bahan baku industri kimia. Penggunaannya menimbulkan risiko tumpahan, pencemaran lingkungan, serta emisi berbahaya saat dibakar

2. Energy



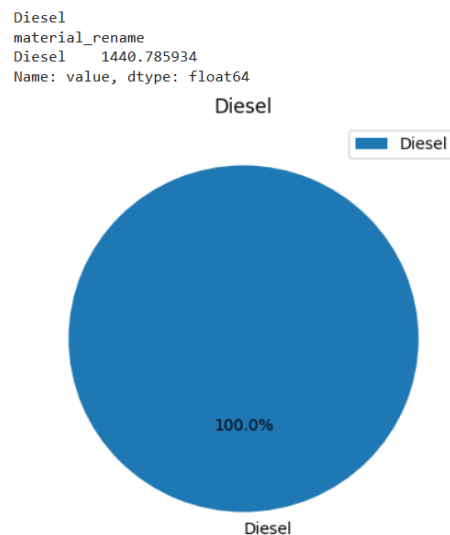
Gambar 4. 25 Kategori Energi Proses Use Stage

- Electricity: 1,850,105 kWh

Electricity (listrik) adalah energi yang digunakan untuk menggerakkan peralatan, mesin, penerangan, dan sistem produksi di berbagai sektor.

Meskipun listrik sendiri bersih saat digunakan, dampak lingkungannya tergantung pada sumber pembangkitannya. Jika listrik dihasilkan dari pembakaran bahan bakar fosil seperti batu bara atau gas alam, maka akan menghasilkan emisi gas rumah kaca seperti karbon dioksida (CO₂) dan polutan lainnya.

3. Diesel

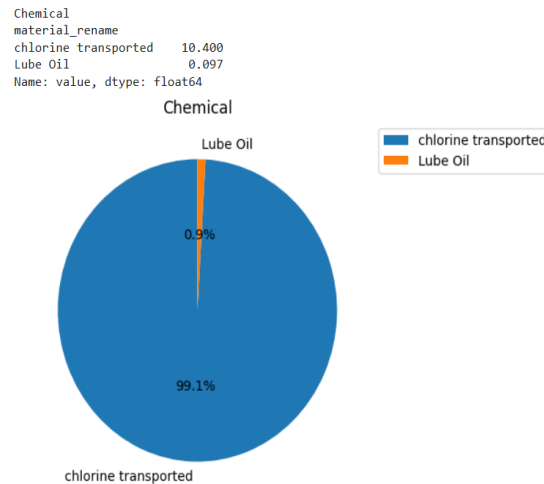


Gambar 4. 26 Kategori Diesel Proses Use Stage

- Diesel: 1440 t

Diesel adalah bahan bakar dari minyak bumi yang digunakan pada kendaraan berat, mesin industri, dan pembangkit listrik. Meski efisien, pembakarannya menghasilkan polusi udara seperti NO₂, CO₂, dan partikulat yang berdampak buruk bagi kesehatan dan lingkungan.

4. Chemical



Gambar 4. 27 Kategori Chemical Proses Use Stage

- Chlorin transported: 10.4 t

Chlorin Transported adalah kegiatan pengangkutan gas atau cairan klorin, biasanya untuk keperluan industri atau pengolahan air. Klorin bersifat beracun dan korosif, sehingga proses transportasinya harus sangat hati-hati untuk mencegah kebocoran yang membahayakan manusia dan lingkungan.

- Lube oil: 0.1 t

Lube Oil (oli pelumas) adalah cairan yang digunakan untuk mengurangi gesekan dan panas pada mesin. Penggunaannya penting untuk menjaga performa dan umur mesin, tetapi jika tidak dikelola dengan benar, limbah oli dapat mencemari tanah dan air.

b. Proses HRSG (*Heat Recovery Steam Generator*)

```
# check locations with HRSG
df_pp_hrsg['location'].unique()

array(['IP Grati', 'IP Cilegon'], dtype=object)
```

Gambar 4. 28 Check Location By HRSG

Dapat dilihat bahwa terdapat 2 tempat yang melakukan unit process HRSG. Untuk mengetahui kenapa unit process ini menghasilkan emisi yang banyak, akan di tracing input pada lokasi-lokasi ini.

	subcategory	value
0	Material	58379.943825
1	Fuel	29189.971913
2	Chemical	9.089000

Gambar 4. 29 Emisi Material Proses HRSG

Dari Lokasi IP Grati dan IP Cilegon, diketahui ada beberapa kategori material yang digunakan sebagai input, di antaranya:

- Fuel: 55,969,580 kl
Fuel (bahan bakar) adalah sumber energi utama yang digunakan untuk pembakaran dalam mesin, kendaraan, dan industri. Jenis bahan bakar bisa berupa bensin, solar, gas alam, dan lainnya. Penggunaan bahan bakar ini menghasilkan emisi gas rumah kaca seperti CO₂ dan polutan lain yang mencemari udara.

- Material: 3,031,746

Material mengacu pada bahan mentah atau setengah jadi yang digunakan dalam proses produksi, seperti logam, plastik, kayu, dan lainnya. Pengelolaan material yang tidak efisien dapat menyebabkan limbah dan emisi tambahan dalam proses industri.

- Chemical: 9.9 t

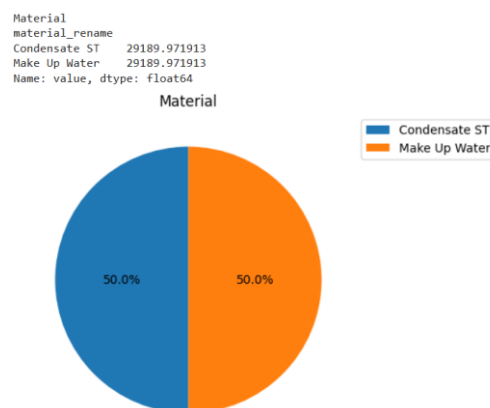
Chemical (bahan kimia) digunakan luas di industri untuk proses produksi, pengawetan, dan pembersihan. Penggunaan dan pembuangan bahan kimia yang tidak terkendali dapat mencemari udara, air, dan tanah serta membahayakan kesehatan manusia dan lingkungan

Kategori-kategori input pada proses HRSG tersebut akan dianalisis lebih lanjut lagi untuk melihat material apa saja yang digunakan, Berikut hasil analisis material dari kategori input :

1. Material

Condensate ST: 2,923,788.47 t

Make Up Water: 107,957.77 t



Gambar 4. 30 Kategori Material Proses HRSG

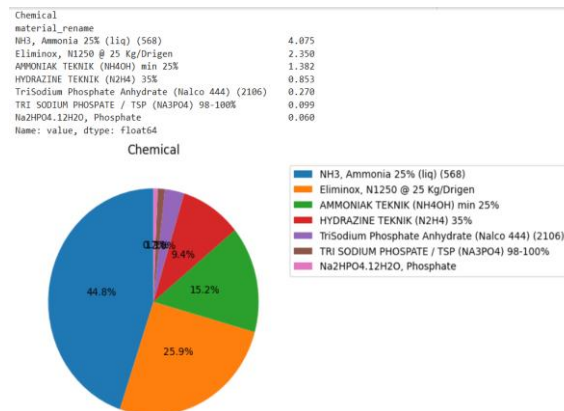
- Condensate ST

Condensate ST (Steam Condensate) sebesar 2,923,788.47 ton adalah air hasil kondensasi uap dalam sistem pembangkit atau industri, yang biasanya dikembalikan ke sistem untuk digunakan kembali sebagai air umpan boiler. Ini membantu efisiensi energi dan mengurangi konsumsi air.

- Make Up Water

Make Up Water sebesar 107,957.77 ton adalah air tambahan yang disuplai untuk menggantikan kehilangan air dari sistem (karena penguapan, kebocoran, atau pembuangan). Air ini penting untuk menjaga keseimbangan dan tekanan dalam sistem uap atau pendingin.

2. Chemical



Gambar 4. 31 Kategori Chemical Proses HRSG

- NH3, Ammonia 25% (liq) (568): 4.08 t

Meskipun digunakan untuk mengontrol pH dan mengurangi NOx, jika terjadi kebocoran atau penguapan, amonia dapat mencemari udara dan menyebabkan iritasi pernapasan serta bau menyengat.

- Eliminox, N1250 @ 25 Kg/Drigen: 2.35 t

Produk ini umumnya digunakan dalam air proses, namun jika terbakar atau menguap, zat kimianya bisa menghasilkan senyawa organik volatil (VOC) yang berkontribusi pada pencemaran udara.

- Ammoniak teknik (NH_4OH) min 25%: 1.38 t

Serupa dengan NH_3 , ammoniak teknik berpotensi melepaskan gas amonia ke udara jika tidak ditangani dengan benar, yang dapat menyebabkan pencemaran udara lokal dan risiko kesehatan.

- Hydrazine teknik (N_2H_4) 35%: 0.85 t

Hydrazine sangat reaktif dan beracun. Bila bocor atau menguap, gas hydrazine dapat mencemari udara dan menimbulkan dampak serius pada kesehatan, termasuk kerusakan paru-paru dan kanker.

- TriSodium Phosphate Anhydrate (Nalco 444) (2106): 0.27 t

Secara umum, zat ini tidak mudah menguap, sehingga tidak langsung menyebabkan emisi udara. Namun, pengelolaan yang buruk bisa menyebabkan dekomposisi kimia yang melepaskan gas atau debu ke udara.

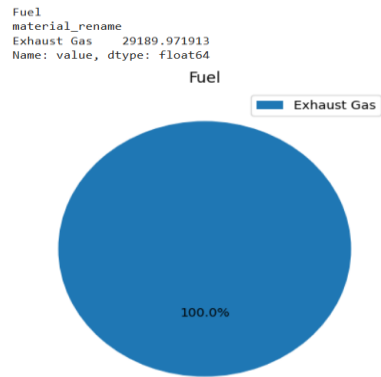
- Tri sodium phosphate / TSP (Na_3PO_4) 98-100%: 0.10 t

Sama seperti Nalco 444, TSP tidak langsung menghasilkan emisi udara, namun penyimpanan atau pembakarannya bisa menghasilkan partikulat halus jika tidak ditangani dengan benar.

- $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, Phosphate: 0.06 t

Risiko emisi udara dari senyawa ini rendah, tetapi partikel debu dari fosfat padat bisa terlepas ke udara dan mengganggu kualitas udara lokal, terutama di area penyimpanan terbuka.

3.Fuel



Gambar 4. 32 Kategori Fuel Proses HRSG

- Exhaust Gas: 55,969,576.71 kCal

Exhaust gas sebesar 55.969.576,71 kCal adalah panas yang masih tersisa dalam gas buang hasil pembakaran di mesin atau pembangkit. Gas ini mengandung polutan seperti CO₂, NO₂, SO₂, dan partikulat yang berdampak pada pencemaran udara. Energi panas tersebut bisa dimanfaatkan kembali, namun jika dibiarkan keluar tanpa pemulihan, tidak hanya membuang energi, tetapi juga meningkatkan emisi yang berbahaya bagi lingkungan dan kesehatan.

4.5 Transformasi Data

4.5.1 Label Encoding

Sebelum dilakukan pemodelan, data harus diproses terlebih dahulu. Salah satu tahap pra-pemrosesan yang dilakukan adalah mengubah data kategorikal menjadi numerik menggunakan label encoding agar data dapat digunakan dalam analisis k-means clustering.

```

import pandas as pd
from sklearn.preprocessing import LabelEncoder
# Mengidentifikasi kolom kategorikal
categorical_cols = df.select_dtypes(include=['object']).columns.tolist()

# Melakukan label encoding pada kolom kategorikal
label_encoders = {}
for col in categorical_cols:
    le = LabelEncoder()
    df[col] = le.fit_transform(df[col])
    label_encoders[col] = le # Menyimpan encoder jika diperlukan untuk decoding di masa mendatang

# Menampilkan informasi setelah label encoding
print("\nInformasi Dataset Setelah Label Encoding:")
print(df.info())

# Menampilkan beberapa baris pertama setelah label encoding
print("\nBeberapa Baris Pertama Setelah Label Encoding:")
print(df.head())

```

Gambar 4. 33 *Code Label Encode*

Kode pada gambar 4.37 bertujuan untuk mempersiapkan data kategorikal agar dapat digunakan dalam analisis lebih lanjut atau dalam pemodelan machine learning, dengan mengubah data yang bertipe kategorikal menjadi nilai numerik melalui proses label encoding, serta memeriksa hasil perubahan tersebut.

```

Informasi Dataset Setelah Label Encoding:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 171 entries, 0 to 170
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   industry              171 non-null   int64
1   sector                171 non-null   int64
2   location              171 non-null   int64
3   unit_process          171 non-null   int64
4   input_output          171 non-null   int64
5   category              171 non-null   int64
6   subcategory           171 non-null   int64
7   material_rename       171 non-null   int64
8   value                 171 non-null   float64
9   unit                  171 non-null   int64
dtypes: float64(1), int64(9)
memory usage: 13.5 KB
None

```

Gambar 4. 34 Informasi Dataset Setelah *Label Encode*

	industry	sector	location	unit_process	input_output	category	subcategory	material_rename	value	unit
0	1	4	9	5	0	1	0	1	172754.23000	0
1	1	4	9	5	0	1	0	13	65.20000	0
2	1	4	9	5	0	1	0	9	1966.29000	0
3	1	4	9	5	0	1	0	12	49.57000	0
4	1	4	9	9	0	1	0	1	601.73749	0

Gambar 4. 35 Sampel Dataset Setelah *Label Encode*

Gambar 4.39 menunjukkan sampel dataset setelah proses label encoding diterapkan pada kolom-kolom kategorikal. Setelah label encoding, setiap kolom yang sebelumnya bertipe object (seperti string atau teks) sekarang menjadi numerik. Misalnya, kolom industri yang awalnya berisi nilai "Powerplant" dan "Oil & Gas Distribution" menjadi berisi nilai 0 dan 1, yang memudahkan model dalam menghitung dan mengolah data.

4.5.2 Feature Selection

Memilih dan menyiapkan fitur yang akan digunakan dalam model Proses ini sering disebut juga sebagai *feature selection* dan *feature engineering*.

```
features = ['sector', 'location', 'unit_process', 'material_rename', 'value']
```

Gambar 4. 36 Kode Pemilihan Fitur

Feature selection(pemilihan fitur) bertujuan untuk memilih fitur yang memiliki pengaruh besar terhadap hasil yang ingin dicapai oleh pemodelan. Dalam dataset yang digunakan, terdapat beberapa kolom atau atribut yang

memiliki potensi untuk memberikan informasi yang penting mengenai karakteristik emisi udara yang dianalisis.

Kolom-kolom yang dipilih sebagai fitur dalam pemodelan atau pengelompokkan adalah:

- a. Sector: Mengelompokkan data berdasarkan sektor, membantu mempermudah pola yang spesifik untuk masing-masing sektor yang dapat memberikan informasi penting untuk mengurangi emisi dan meningkatkan efisiensi material.
- b. Location (Lokasi): Lokasi sangat penting untuk memahami bagaimana faktor eksternal mempengaruhi emisi dan penggunaan material dalam proses industri.
- c. Unit Process: Memahami unit process memungkinkan kita untuk mengidentifikasi tahapan proses yang paling berkontribusi terhadap emisi dan penggunaan material, serta mengoptimalkan proses tersebut untuk meningkatkan efisiensi dan mengurangi dampak lingkungan.
- d. Material Rename : Material memiliki pengaruh langsung terhadap kinerja operasional dan lingkungan dikarenakan *attribute* tersebut berisi data emisi yang dihasilkan sehingga sangat *relevan* atau sesuai dengan tujuan analisis.
- e. Value : Value berisikan informasi mengenai jumlah material atau emisi yang digunakan dari masing – masing sektor berdasarkan unit proses oleh karena itu atribut value sangat *relevan* (Sesuai).

Pemilihan fitur yang sesuai merupakan langkah krusial dalam pra-pemrosesan data. Dengan memilih atribut yang tepat, kita dapat memastikan bahwa model yang dibangun akan memiliki basis data yang kuat dan mampu memberikan prediksi yang akurat terkait pengelompokkan sumber emisi udara berdasarkan karakteristiknya menggunakan metode k-means clustering.

4.5.3 Standarisasi data

Standarisasi data adalah bagian dari tahapan transformasi/pra-pemrosesan data yang harus dilakukan sebelum tahap modeling. Langkah ini memastikan bahwa data memiliki skala yang seragam, yang sangat penting untuk algoritma machine learning yang sensitif terhadap skala data, seperti K-Means Clustering.

```
# Normalisasi data
scaler = StandardScaler()
df_scaled = scaler.fit_transform(df_selected)
```

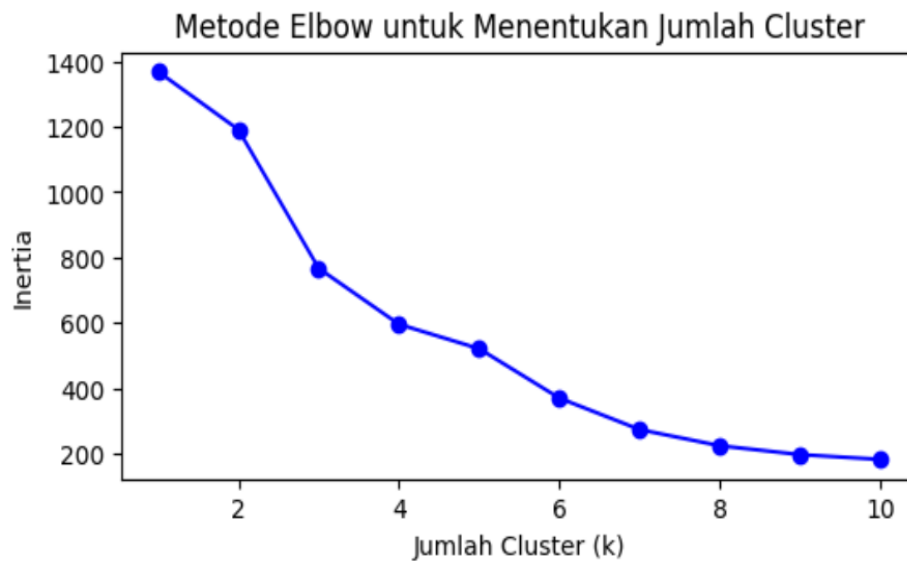
Gambar 4. 37 Kode Normalisasi Data

Pada gambar 4.41 menampilkan kode analisis yang digunakan untuk menormalkan data numerik (sudah dipilih sebelumnya). *Standardscaler* adalah salah satu metode yang digunakan untuk normalisasi atau standarisasi data numerik agar setiap fitur berada pada skala yang seragam dengan rata-rata 0 dan standar deviasi 1. Hal ini akan memastikan bahwa algoritma machine learning yang digunakan nantinya tidak terpengaruh oleh perbedaan skala antara fitur-fitur yang ada.

4.6 Modeling (Data Mining)

4.6.1 Penentuan Jumlah Cluster

Penentuan jumlah cluster menggunakan metode elbow, metode elbow adalah teknik yang digunakan dalam analisis cluster untuk menentukan jumlah kelompok (cluster) yang paling tepat untuk suatu dataset. Tujuannya adalah untuk menemukan titik optimal dimana penambahan cluster tidak lagi memberikan peningkatan signifikan dalam kualitas model.



Gambar 4. 38 Visualisasi Metode *Elbow*

Analisis metode elbow menunjukkan bahwa jumlah cluster optimal untuk data penelitian ini adalah 3. Hal ini terlihat dari penurunan WCSS (*Within-cluster sum of squares*) yang semakin melambat setelah jumlah cluster mencapai 3. Dengan demikian, data penelitian dibagi menjadi 3 kelompok yang berbeda. Hasil visualnya menunjukkan bahwa terdapat tiga segmen pola perilaku yang berbeda-beda. Pemilihan 3 cluster ini dianggap optimal karena dapat mengidentifikasi pola yang jelas dalam data tanpa mengorbankan interprebilitas hasil.

4.6.2 Algoritma K-Means Clustering

Teknik *clustering* merupakan suatu pendekatan analisis data yang bertujuan untuk mengelompokkan entitas data secara otomatis berdasarkan kemiripan karakteristiknya, tanpa adanya label kelas yang telah ditentukan sebelumnya. *Clustering*, seringkali diidentifikasi sebagai metode *unsupervised learning*, memberikan kemampuan untuk mengklasifikasikan data yang belum memiliki label kelas yang jelas atau terdefinisi. Metode ini menjadi esensial dalam konteks

analisis data kompleks, di mana informasi tentang struktur kelas tidak diketahui atau sulit ditentukan. Proses clustering melibatkan identifikasi pola-pola intrinsik dalam data, dengan tujuan membentuk kelompok-kelompok homogen yang memiliki tingkat kemiripan tinggi di antara anggotanya.

Adapun pada penelitian ini, saya menggunakan klasterisasi metode k-means. Algoritma ini bekerja dengan menjalankan perintah banyak kelas klaster (k) yang diinginkan. Selain itu, k-means bekerja dengan menunggu jumlah pengulangan eksekusi perintah, yang dikenal dengan istilah iterasi, maksimum dan sebuah nilai konstanta yang digunakan untuk penentuan kapan iterasi dihentikan.

Setelah jumlah cluster ditentukan, langkah selanjutnya adalah menerapkan algoritma K-Means clustering pada data yang sudah dinormalisasi. K-Means clustering bekerja dengan cara mengelompokkan data ke dalam k cluster berdasarkan kedekatannya dengan pusat cluster (centroid). Setiap data akan dikelompokkan ke dalam cluster yang memiliki jarak terdekat dengan centroid tersebut. Berikut penggunaan algoritma k-means pada penelitian ini dengan nilai $K=3$:

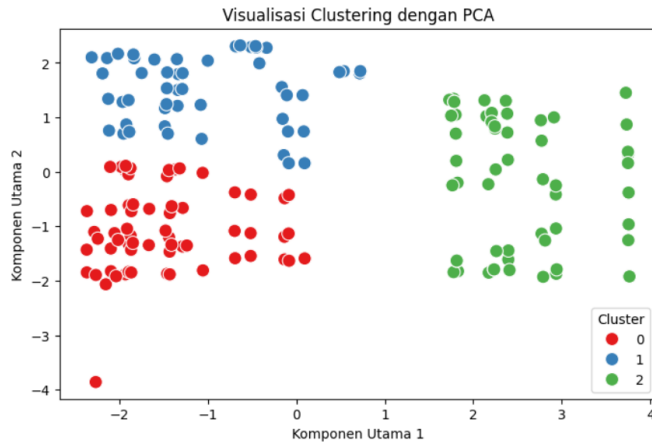
```
# Clustering menggunakan K-Means (misalnya, 3 cluster)
kmeans = KMeans(n_clusters=3, random_state=42)
clusters = kmeans.fit_predict(df_scaled)

# Menambahkan hasil clustering ke data asli
df['Cluster'] = clusters

# Analisis hasil clustering
# Visualisasi PCA untuk 2D plot
pca = PCA(n_components=2)
df_pca = pca.fit_transform(df_scaled)

plt.figure(figsize=(8, 5))
sns.scatterplot(x=df_pca[:, 0], y=df_pca[:, 1], hue=df['Cluster'], palette='Set1', s=100)
plt.title('Visualisasi Clustering dengan PCA')
plt.xlabel('Komponen Utama 1')
plt.ylabel('Komponen Utama 2')
plt.legend(title='Cluster')
plt.show()
```

Gambar 4. 39 *Code Metode K-Means Clustering*



Gambar 4. 40 Visualisasi *Clustering PCA*

Pada gambar 4.44 menampilkan visualisasi clustering dengan PCA, terdapat 3 cluster optimal yang telah di tentukan nilai K sebelumnya menggunakan metode elbow, visualisasi tersebut menggambarkan cluster 0 persebarannya di tandai dengan warna merah, cluster 1 berwarna biru dan cluster 2 berwarna hijau.

Setelah dilakukan pengolahan data dan penerapan K-Means Clustering, berikut adalah persentase per cluster yang dihasilkan dari analisis:

Tabel 4. 1 Jumlah Persentase Data *Cluster*

Cluster	Jumlah Data	Persentase
0	63	37%
1	50	29%
2	58	34%

Tabel 4.2 menunjukkan jumlah data dan persentase data yang termasuk dalam masing-masing klaster setelah dilakukan analisis K-Means Clustering dengan jumlah klaster (K) = 3.

- a. Cluster 0: Memiliki 63 data, yang berarti sekitar 37% dari total data berada dalam klaster ini.
- b. Cluster 1: Memiliki 50 data, yang berarti sekitar 29% dari total data berada dalam klaster ini.
- c. Cluster 2: Memiliki 58 data, yang berarti sekitar 34% dari total data berada dalam klaster ini.

Persentase ini memberikan gambaran mengenai distribusi data dalam setiap klaster, yang membantu dalam memahami seberapa besar proporsi masing-masing klaster dalam keseluruhan dataset.

4.6.3 Hasil Analisis Clustering dengan Nilai $K=3$

Analisis Cluster Berdasarkan hasil analisis menggunakan metode K-Means clustering, data emisi udara berdasarkan karakteristiknya berhasil dikelompokkan ke dalam tiga cluster dengan karakteristik emisi yang berbeda-beda terkait dengan karakteristik dari masing-masing sumber emisi tersebut .

analisis clustering dengan metode K-Means terhadap data output berupa emisi dari industri Power Plant dan Oil & Gas Distribution menghasilkan tiga klaster utama yang memiliki karakteristik berbeda berdasarkan kombinasi sektor, lokasi, unit proses, jenis emisi, dan besaran emisi (value). Adapun interpretasi setiap cluster dijelaskan sebagai berikut:

Hasil Clustering (Rata-rata Fitur per Cluster):

	industry	sector	location	unit_process	input_output	category	subcategory	material	material_rename	value	unit
Cluster											
0	0.00	0.809524	5.761905	11.873016	0.0	1.000000	0.0	13.047619	4.174603	6.406547e+07	0.0
1	0.08	1.160000	7.940000	10.860000	0.0	1.000000	0.0	42.100000	14.540000	2.241744e+02	0.0
2	1.00	4.068966	9.913793	3.844828	0.0	0.844828	0.0	22.431034	8.344828	1.103890e+06	0.0

Gambar 4. 41 Rata-rata Hasil *Cluster*

Berdasarkan pengelompokkan didapatkan hasil cluster dengan karakteristik dari masing-masing kelompok yaitu :

- Cluster 0 (High Air Emission) merupakan kelompok emisi udara sangat tinggi dengan rata-rata nilai emisi yaitu mencapai 64,065,473.17 ton, berasal dari industri Oil & Gas Distribution. Emisi dominan berasal dari proses pembakaran bahan bakar fosil, khususnya pada sektor Injection Terminal (IT) dan proses Use Stage. Jenis emisi utama adalah Carbon Dioxide (CO₂) yang sangat besar volumenya, sehingga cluster ini memberikan dampak besar terhadap pemanasan global dan kualitas udara.
- Cluster 1 (Low Air Emission) adalah kelompok emisi udara rendah dengan rata-rata nilai emisi hanya mencapai 224,17 ton, yang terutama berasal dari proses penguapan bahan bakar atau pelarut kimia di sektor Fuel Terminal(FT) dan Storage Tank. Emisi utama berupa Volatile Organic Compounds (VOC) yang mudah menguap dan jumlahnya jauh lebih kecil dibanding cluster lain. Karena bersumber dari proses non-pembakaran dan volume emisinya kecil, dampak lingkungan cluster ini tergolong rendah.

- c. Cluster 2 (Moderate Air Emission) mencakup sumber emisi tingkat sedang dengan rata-rata nilai emisi 1,103,889.96 ton, berasal dari industri pembangkit listrik, terutama sektor PLTG/PLTGU. Emisi terbesar dihasilkan dari proses pembakaran di unit HRSG, dengan Nitrogen Dioxide (NO₂) sebagai jenis emisi dominan. Total emisi cluster ini lebih besar dari cluster rendah namun jauh lebih kecil dibanding cluster tinggi, sehingga dampaknya berada di tingkat menengah terhadap lingkungan.

4.6.4 Perbandingan Rata-rata Cluster

Berdasarkan hasil analisis clustering, diperoleh rata-rata persentase untuk masing-masing cluster sebagai berikut:

Tabel Persentase per Klaster:

Cluster	Jumlah Data	Persentase
0	63	36.842105
1	50	29.239766
2	58	33.918129

Gambar 4. 42 Tabel Rata-rata Jumlah *Cluster*

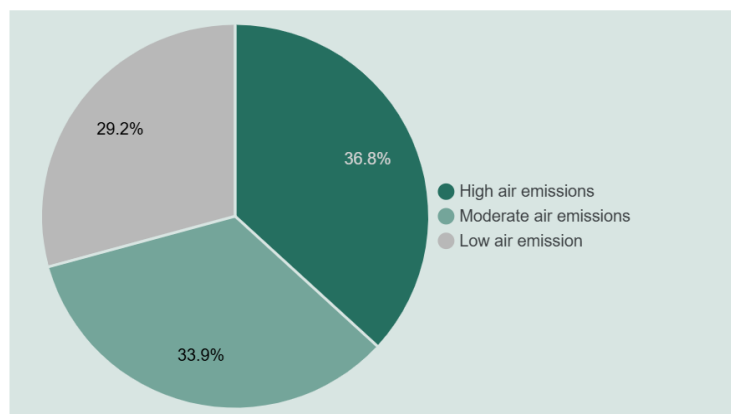
- a. Cluster 0 merupakan kelompok dengan jumlah data terbanyak, mencakup 63 unit atau sekitar 36,84% dari total keseluruhan. Klaster ini didominasi oleh unit-unit yang cenderung menghasilkan output dalam skala besar, dengan proses operasional yang kompleks namun variasi material input yang terbatas. Unit-unit dalam klaster ini memainkan peran penting dalam sistem distribusi energi perusahaan dan perlu dikawal dari sisi efisiensi serta pengelolaan dampak lingkungan yang terfokus.
- b. Cluster 1 adalah klaster dengan jumlah data paling sedikit, yakni 50 unit atau 29,24% dari total. Unit-unit dalam klaster ini menunjukkan karakteristik unik, dengan tingkat output yang sangat kecil tetapi jenis material yang digunakan

sangat beragam. Hal ini mengindikasikan bahwa klaster ini kemungkinan mencakup unit-unit beroperasi skala kecil. Meskipun kontribusinya terhadap total emisi relatif rendah, banyaknya variasi material yang digunakan tetap memerlukan perhatian, terutama dalam pengelolaan limbah, dokumentasi bahan berbahaya, dan efisiensi operasional mikro.

- c. Cluster 2 mencakup 58 unit data atau sekitar 33,92% dari keseluruhan, dan mewakili kelompok unit yang beroperasi sebagai fasilitas industri inti seperti kilang minyak. Klaster ini dicirikan oleh penggunaan material dalam variasi dan volume sedang, dengan output emisi yang juga berada pada tingkat menengah. Dengan lokasi yang cenderung strategis dan orientasi produksi yang aktif, unit-unit dalam Cluster 2 merupakan bagian vital dari rantai produksi energi perusahaan Ailesh. Karena berada dalam kelas operasional aktif, unit-unit ini perlu menjadi fokus dalam penerapan teknologi pengurangan emisi dan efisiensi energi secara sistematis.

4.5 Hasil Analisis Visualisasi

1. Visualisasi Data Berdasarkan Cluster

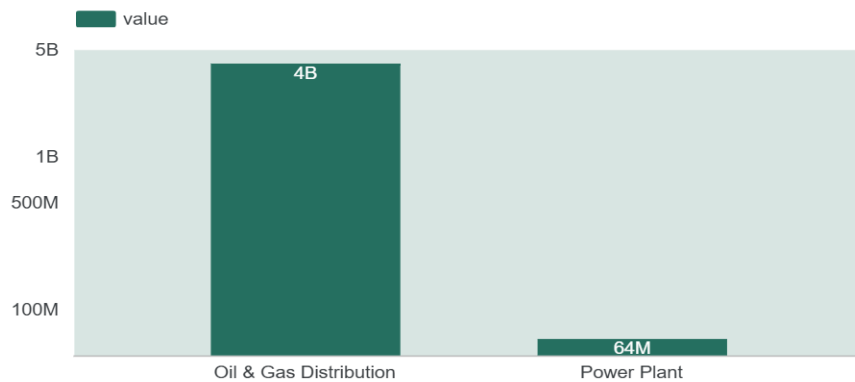


Gambar 4. 43 Visualisasi Data *Cluster*

Visualisasi pada gambar 4.47 memberikan informasi penting mengenai pengelompokan data sumber emisi udara yang termasuk dalam masing-masing kategori berdasarkan tingkat emisi yang di hasilkan. Ketiga kategori tersebut adalah :

- Cluster 0 (High air emission) merupakan kelompok dengan jumlah data terbanyak, mencakup sekitar 36,8% dari total data.
- Cluster 1 (Low air emission) merupakan kelompok dengan jumlah data paling sedikit, yaitu dengan persentase 29,2% dari total data.
- Cluster 2 (Moderate air emission) merupakan kelompok dengan jumlah data yang lebih banyak dibandingkan dengan cluster 1 , yaitu dengan persentase 33,9 % dari total data.

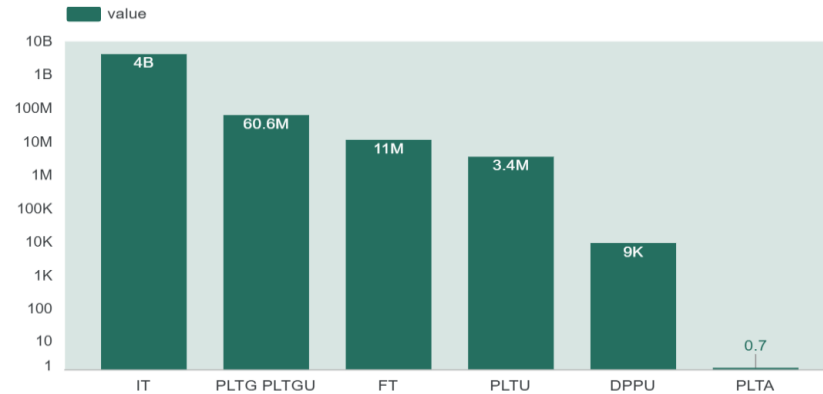
2. Visualisasi Data Berdasarkan Industri



Gambar 4. 44 Visualisasi Data Emisi Industri

Berdasarkan perhitungan nilai emisi yang telah dilakukan pada industri oil & gas distribution dan power plant, ditemukan emisi sebanyak 4,036,136,000 ton pada industri oil & gas distribution sedangkan pada industri power plant ditemukan emisi sebanyak 64,065,400 ton. dapat disimpulkan bahwa industri oil & gas distribution menghasilkan emisi udara jauh lebih banyak dari industri power plant.

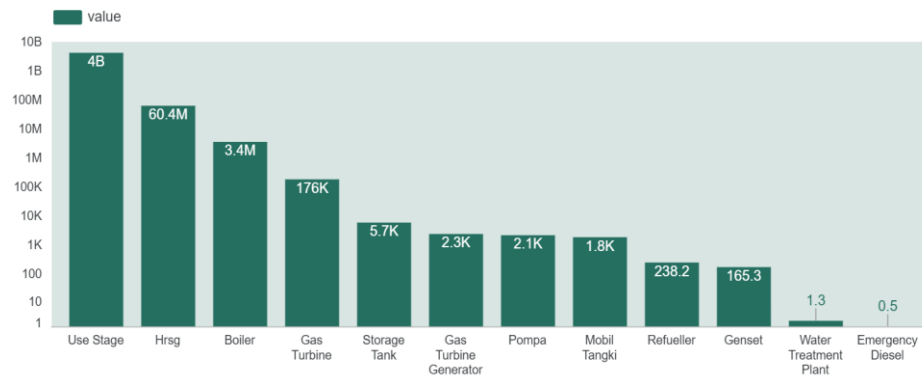
3. Visualisasi Data Jumlah Emisi Berdasarkan Sektor



Gambar 4. 45 Visualisasi Data Emisi Sektor

Visualisasi pada gambar 4.49 menampilkan informasi mengenai sektor yang menghasilkan emisi paling banyak, berdasarkan informasi dari visualisasi sumber emisi berasal dari sektor IT, PLTG/PLTGU, FT, PLTU, DPPU, dan PLTA. Dari sektor – sektor tersebut dapat dilihat bahwa penghasil emisi yang berada pada urutan pertama yaitu sektor IT dengan total nilai mencapai 4,025,105,296.04 ton.

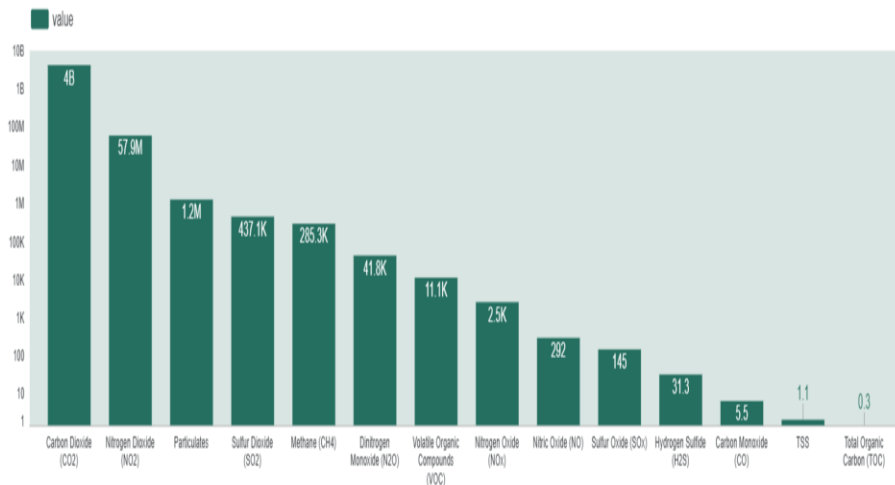
4. Visualisasi Unit Proses Berdasarkan Jumlah Emisi



Gambar 4. 46 Visualisasi Emisi Unit Proses

Visualisasi unit proses pada gambar 4.50 memberikan informasi mengenai peralatan utama yang terlibat dalam produksi atau distribusi energi, pengolahan bahan bakar, atau proses lain di dalam industri *power plant* dan *oil & gas distribution*. Dapat dilihat pada visualisasi menunjukkan bahwa pada proses Use Stage menghasilkan emisi paling banyak yaitu mencapai 4,036,124,400.02 ton, selanjutnya proses HRSG menjadi penghasil emisi paling banyak kedua yaitu mencapai 60,413,451.56 ton.

5. Visualisasi Berdasarkan Jenis Emisi Udara yang Dihasilkan

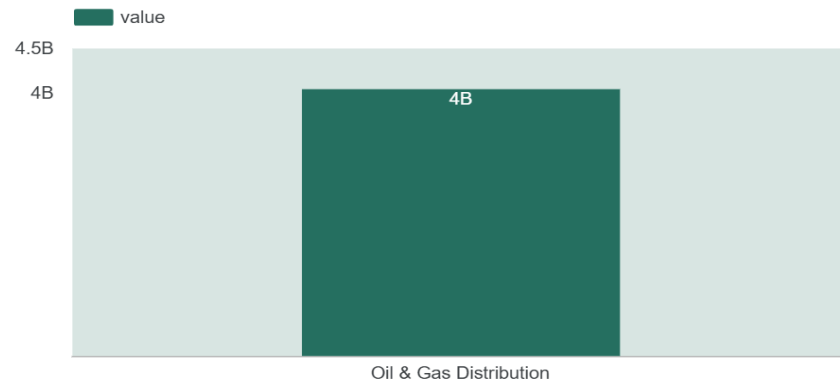


Gambar 4. 47 Visualisasi Jenis Emisi

Visualisasi tersebut menampilkan jenis emisi apa saja yang dihasilkan oleh tiap industri, berdasarkan visualisasi dapat dilihat bahwa Carbon Dioxide (CO₂) menjadi emisi yang paling banyak dihasilkan dari kedua industri tersebut dengan persentase nilai mencapai 98.5 %.

6. Visualisasi Cluster 0 (*High air emission*)

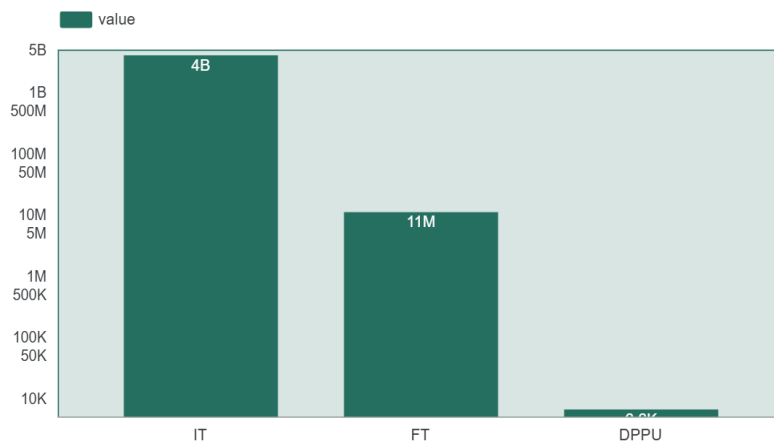
a. Visualisasi *Cluster High Air Emission* Berdasarkan Industri



Gambar 4. 48 Visualisasi Industri *Cluster High Air Emission*

Data visualisasi tersebut menunjukkan bahwa cluster 0 (*High air emission*) merupakan kelompok sumber emisi yang berada pada industri *Oil & Gas Distribution* dengan total nilai mencapai 4,036,124,178.69 ton.

b. Visualisasi *Cluster High Air Emission* Berdasarkan Sektor

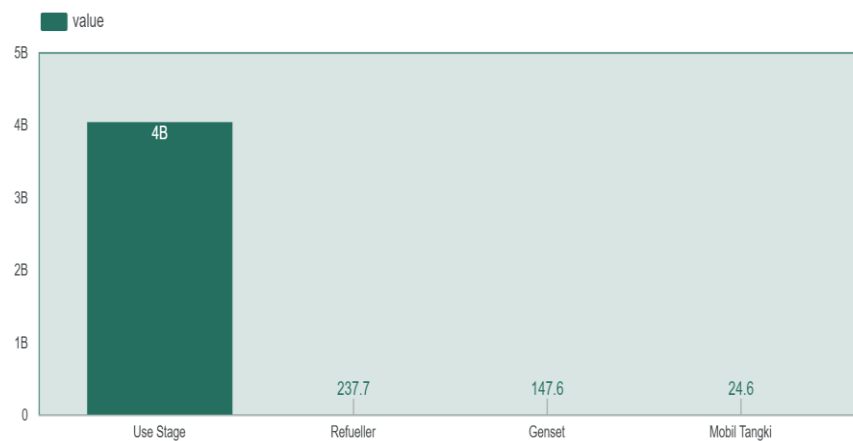


Gambar 4. 49 Visualisasi Sektor *Cluster High Air Emission*

Visualisasi pada gambar 4.53 menampilkan data sumber emisi dari cluster 0 berdasarkan sektor yaitu memberikan informasi mengenai jumlah emisi berdasarkan sektor.

Sektor IT menjadi sektor penghasil emisi terbesar dengan total nilai emisi 4,025,105,147.3 ton , dibandingkan dengan sektor FT total nilai emisi yaitu 11,013,101.16 ton serta sektor DPPU dengan total nilai emisi yang paling kecil sejumlah 6,561.43 ton, termasuk ke dalam cluster 0.

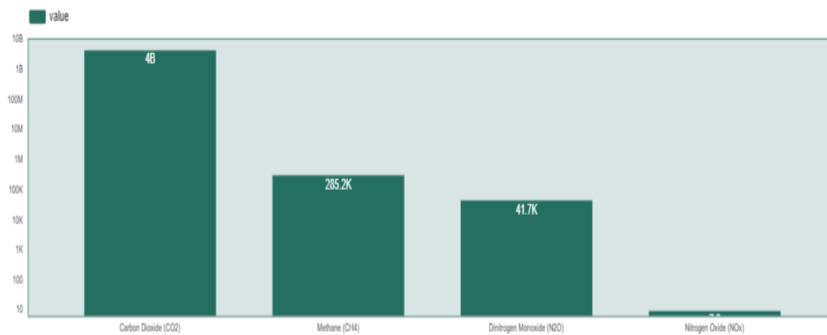
c. Visualisasi *Cluster High Air Emission* Berdasarkan Unit Proses



Gambar 4. 50 Visualisasi Unit Proses *Cluster High Air Emission*

Visualisasi cluster 0 berdasarkan unit proses memberikan informasi jumlah emisi pada unit proses, pada visualisasi tersebut terdapat beberapa proses diantaranya proses Use Stage yang menjadi proses penghasil emisi paling banyak yaitu mencapai 4,03,123,811.62 ton, dibandingkan dengan proses HRSG, Refueller, Genset, dan Mobil Tangki.

d. Visualisasi *Cluster High Air Emission* Berdasarkan Jenis Emisi Udara

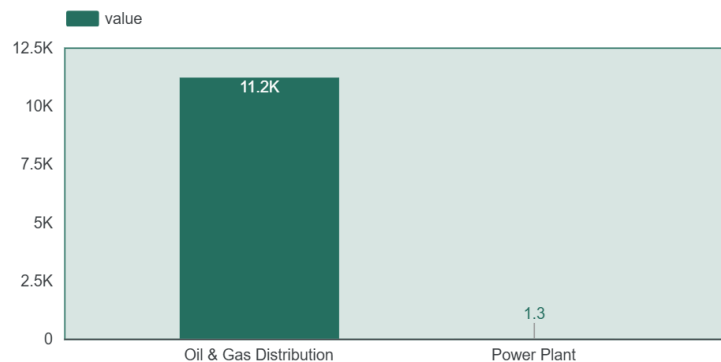


Gambar 4. 51 Visualisasi Jenis Emisi *Cluster High Air Emission*

Berdasarkan visualisasi yang ditampilkan pada gambar 4.55 memberikan informasi jenis emisi udara yang termasuk kedalam cluster 0 (*High air emission*), terdapat beberapa jenis emisi diantaranya adalah Carbon Dioxide (CO₂) merupakan menjadi sumber emisi yang paling banyak pada cluster 0, dibandingkan dengan Methane (CH₄), Dinitrogen Monoxide (N₂O), dan Ammonia (NH₃).

7. Visualisasi *Cluster 1 (Low air Emission)*

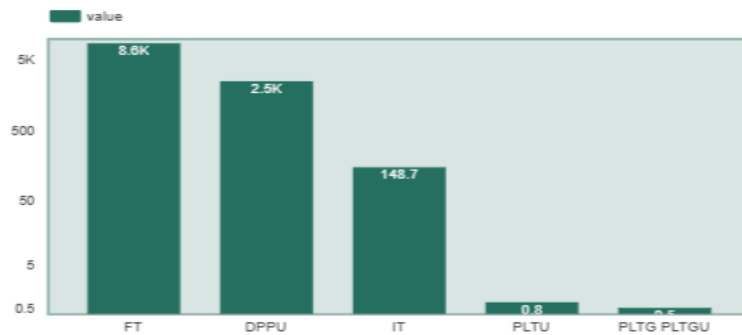
a. Visualisasi *Cluster Low Air Emission* Berdasarkan Industri



Gambar 4. 52 Visualisasi Industri *Cluster Low Air Emission*

Berdasarkan visualisasi pada gambar 4.56 menunjukkan bahwa cluster 1 (*Low air emission*) merupakan kelompok sumber emisi rendah yang berada di antara industri *oil & gas distribution* dan *power plant*. Berdasarkan data industri *oil & gas distribution* lebih dominan daripada industri *power plant*.

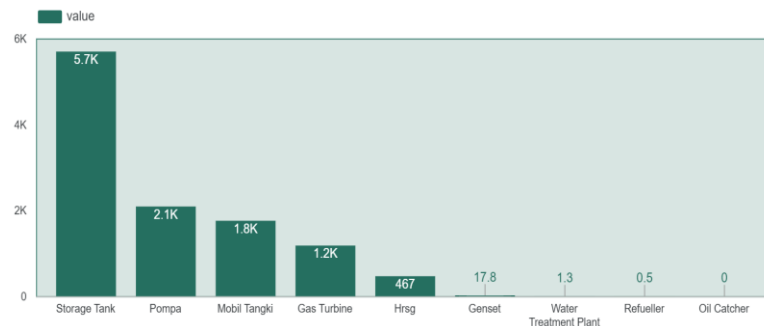
b. Visualisasi *Cluster Low Air Emission* Berdasarkan Sektor



Gambar 4. 53 Visualisasi Sektor *Cluster Low Air Emission*

Visualisasi tersebut menampilkan data sumber emisi dari cluster 1 yang dilihat berdasarkan sektor yaitu memberikan informasi bahwa sektor FT dengan total nilai emisi 8,570.83 ton , sektor DPPU total nilai emisi yaitu 2,487.81 ton, sektor IT dengan total nilai emisi sejumlah 148.73 ton, dan sektor PLTU 0.83 ton, yang terakhir adalah sektor PLTG/PLTGU 0.52 ton.

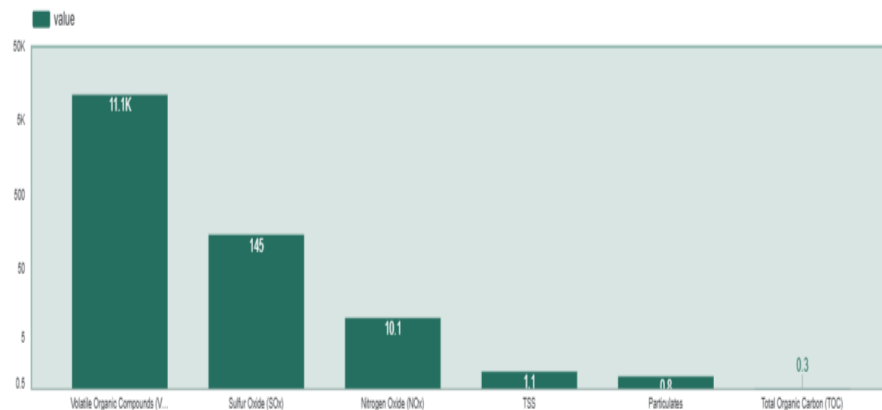
c. Visualisasi *Cluster Low Air Emission* Berdasarkan Unit Proses



Gambar 4. 54 Visualisasi Unit Proses *Cluster Low Air Emission*

Berdasarkan visualisasi pada gambar 4.58 menunjukkan data unit proses yang termasuk kedalam kelompok emisi udara rendah pada cluster 1 diantaranya proses Storage Tank dengan total nilai emisi 5,695.13 ton, proses tersebut menjadi unit proses penghasil emisi paling banyak dibandingkan dengan proses yang lainnya.

d. Visualisasi *Cluster Low Air Emission* Berdasarkan Jenis Emisi

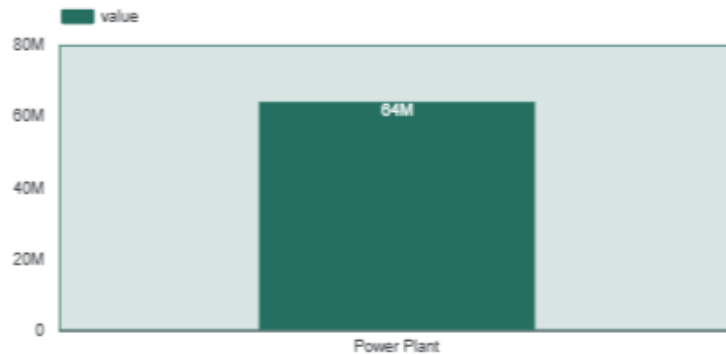


Gambar 4. 55 Visualisasi Jenis Emisi *Cluster Low Air Emission*

Berdasarkan visualisasi gambar 4.59 memberikan informasi jenis emisi udara yang termasuk kedalam cluster 1 (Low air emission), terdapat beberapa jenis emisi diantaranya adalah Volatile Organic Compound (VOC), Sulfur Oxide (Sox), Nitrogen Oxide (Nox), TSS, Particulates dan Total Organic Carbon (TOC). Dapat dilihat dari visualisasi menunjukkan bahwa Volatile Organic Compound (VOC) lebih dominan dibandingkan dengan emisi yang lain.

8. Visualisasi Cluster 2 (*Moderate air Emission*)

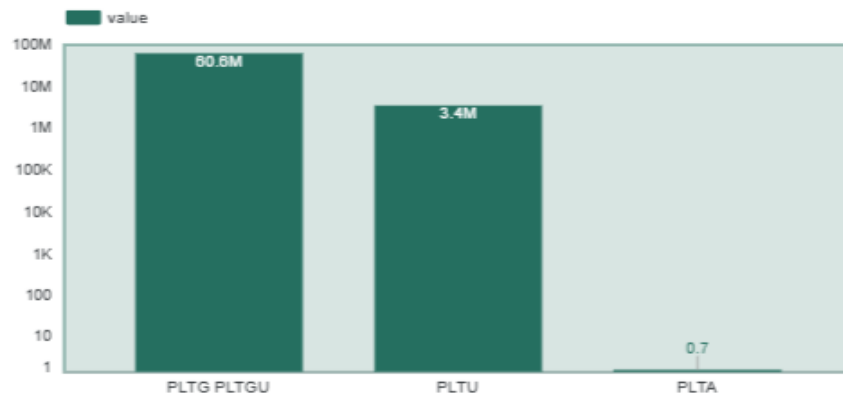
a. Visualisasi *Cluster Moderate Air Emission* Berdasarkan Industri



Gambar 4. 56 Visualisasi Industri *Cluster Moderate Air Emission*

Visualisasi pada gambar 4.60 tersebut menampilkan data industri dari cluster 2 (*Moderate air emission*). *Cluster moderate Air Emission* merupakan kelompok sumber emisi yang lebih tinggi dibandingkan dengan cluster 1. Berdasarkan data yang ditampilkan dalam bentuk visualisasi, cluster 2 berada pada industri *Power Plant* dengan total nilai mencapai 64,025,617.55 ton.

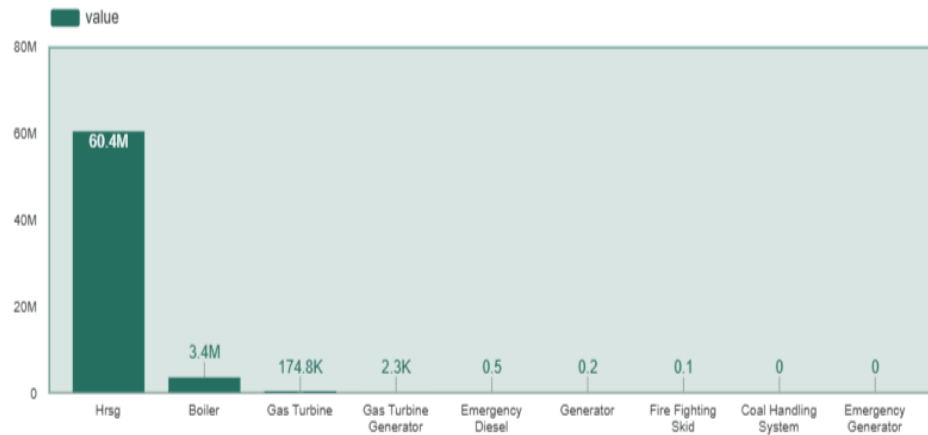
b. Visualisasi *Cluster Moderate Air Emission* Berdasarkan Sektor



Gambar 4. 57 Visualisasi Sektor *Cluster Moderate Air Emission*

Visualisasi berdasarkan gambar 4.61 menampilkan data sumber emisi dari cluster 2 yang dilihat berdasarkan sektor, data tersebut memberikan informasi pada *cluster moderate air emission* sektor PLTG/PLTGU dengan total nilai emisi 60,590,123.37 ton, sektor ini menjadi penghasil emisi paling besar dibandingkan dengan sektor PLTU dengan total nilai emisi yaitu 3,435,494.5 ton, dan sektor PLTA dengan total nilai emisi paling rendah sejumlah 0.7 ton.

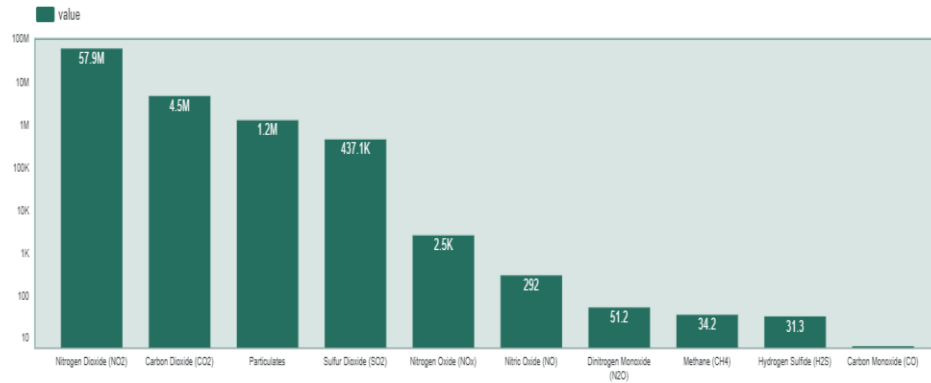
c. Visualisasi *Cluster Moderate Air Emission* Berdasarkan Unit Proses



Gambar 4. 58 Visualisasi Unit Proses *Cluster Moderate Air Emission*

Berdasarkan visualisasi menunjukkan bahwa pada cluster 1 terdapat beberapa proses diantaranya proses HRSG menjadi proses yang menghasilkan emisi paling banyak yaitu mencapai 60,412,984.56 ton, dibandingkan dengan proses Boiler, Gas Turbine, Gas Turbine Generator, Emergency Diesel, Generator, Fire Fighting Skid, Coal Handling System, dan Emergency Generator.

d. Visualisasi *Cluster Moderate Air Emission* Berdasarkan Jenis Emisi



Gambar 4. 59 Visualisasi Jenis Emisi *Cluster Moderate Air Emission*

Berdasarkan visualisasi yang ditampilkan memberikan informasi jenis emisi udara yang termasuk kedalam cluster 2 (Moderate air emission), terdapat beberapa jenis emisi diantaranya adalah Nitrogen Dioxide (NO₂), Carbon Dioxide (CO₂), Particulates, Sulfur Dioxide (SO₂), Nitrogen Oxide (Nox), Nitric Oxide (NO), Dinitrogen Monoxide (N₂O), Methane (CH₄), Hydrogen Sulfide (H₂S), dan Carbon Monoxide (CO). Dapat dilihat dari visualisasi pada *cluster moderate air emission* Nitrogen Dioxide (NO₂) menjadi emisi yang lebih dominan dibandingkan dengan emisi yang lain.