Exploratory Data Analysis for Fuel Emissions

Proponent: Engr. Matthew David M. Loquinerio, REE

Questions to Answer:

1. How much emissions of Fuel Methane and Fuel Nitrous Oxide is emitted from 2010 to 2022?
2. Fuel Emissions of CH4 and N2O fuel by general fuel type.

For this EDA Project I decided to conduct data analyst practice for fuel emission by general fuel type for the timeline starting in 2010 up to year 2022. The CSV file is acquired from Kaggle an open data science forum website.

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib inline
```

Data Reading Process

```python
data = pd.read_csv('data/emissions_by_fuel.csv', encoding='cp1252')
```

```python
data.shape
```

```
(300397, 10)
```

Data Cleaning Process

```python
cols = ['FRS Id', 'Primary NAICS Code', 'Industry Type (subparts)', 'Unit Name', 'Other Fuel Name', 'Blend Fuel Name']
```

```python
data.drop(cols, axis=1, inplace=True)
```

```python
data.rename(columns={
    'Industry Type (sectors)' : 'Industry Type'
}, inplace=True)
```

The file contains 300,397 rows and 10 columns of data. Understanding the dataset some fields are not in the proper data types for transformation. It is substantial for a data analyst to clean the data set it includes: proper usage of data types, replacing unwanted string values, choosing columns that are relevant for analysis, removing duplicated rows, and dropping null subsets.

Exploratory Data Analysis for Fuel Emissions
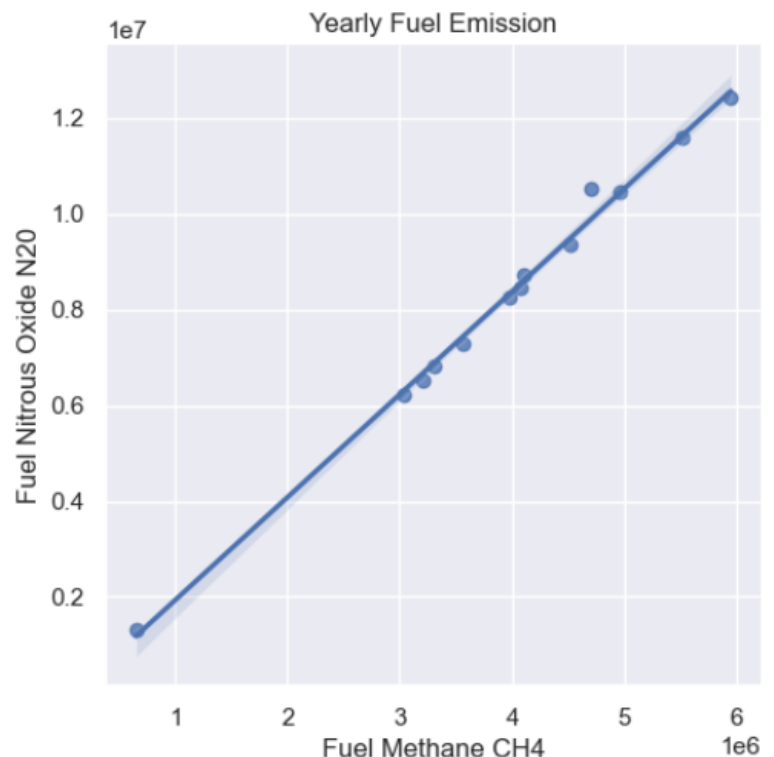
Proponent: Engr. Matthew David M. Loquinerio, REE

Fuel Emissions from 2010 to 2022
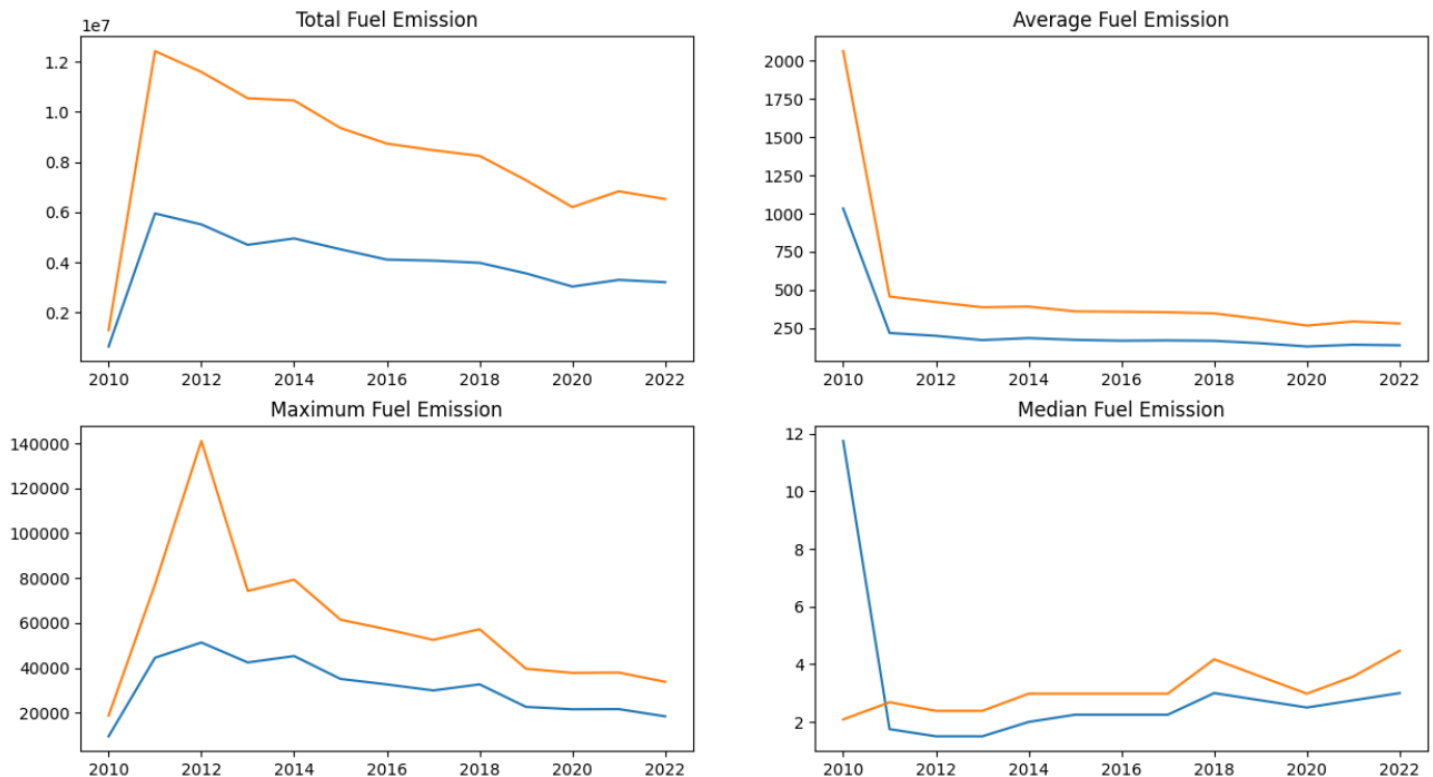
Fuel Emission from 2010 to Year 2022: Total, Average, Maximum, and Median

From 2010 to 2022 we can observe the intermittent decrease on fuel emission. A sudden up spike in fuel usage has been recorded from the start of the decade to its first year in 2011. Until then fuel emission has been reduced and has not been experience sudden uptrend.
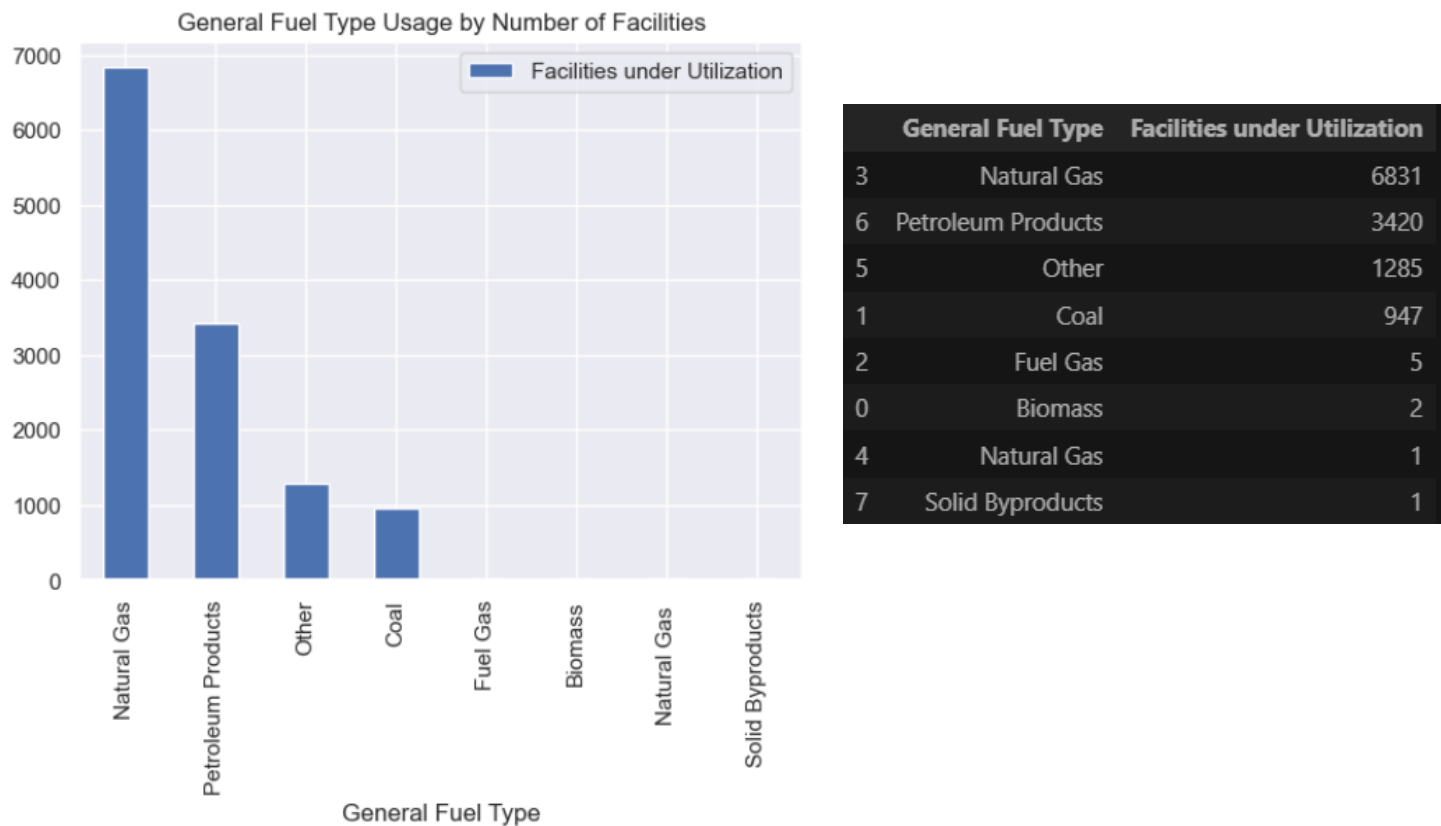
Using bivariate method, we can clearly see that there is not much spread on both quantities of Fuel Methane CH4 and Fuel Nitrous Oxide N20. With that the sample size has low variance. Thus, we can expect a steady trend for few more years.

Exploratory Data Analysis for Fuel Emissions

Proponent: Engr. Matthew David M. Loquinerio, REE

Fuel Emission of CH4 and N20 by General Fuel Type

General Fuel Type Usage by Number of Facilities

|	General Fuel Type	Facilities under Utilization
3	Natural Gas	6831
6	Petroleum Products	3420
5	Other	1285
1	Coal	947
2	Fuel Gas	5
0	Biomass	2
4	Natural Gas	1
7	Solid Byproducts	1

General Fuel Type Ranking and Total Facilities

By summing all the facilities that utilizes a specific fuel type, we deduced the most used fuel type in more than a decade. Ranking them as to show the numbers of facilities with the fuel type of utilization. Natural gas has been growing in numbers constantly developing a total of 6831 facilities under its use. While, petroleum products came in at second for 3420 facilities and coal with 947 facilities. Others as a category can be grouped among the other low utilization fuel types marking them as general outliers for the dataset.

**Exploratory Data Analysis for Fuel Emissions**

Proponent: Engr. Matthew David M. Loquinerio, REE
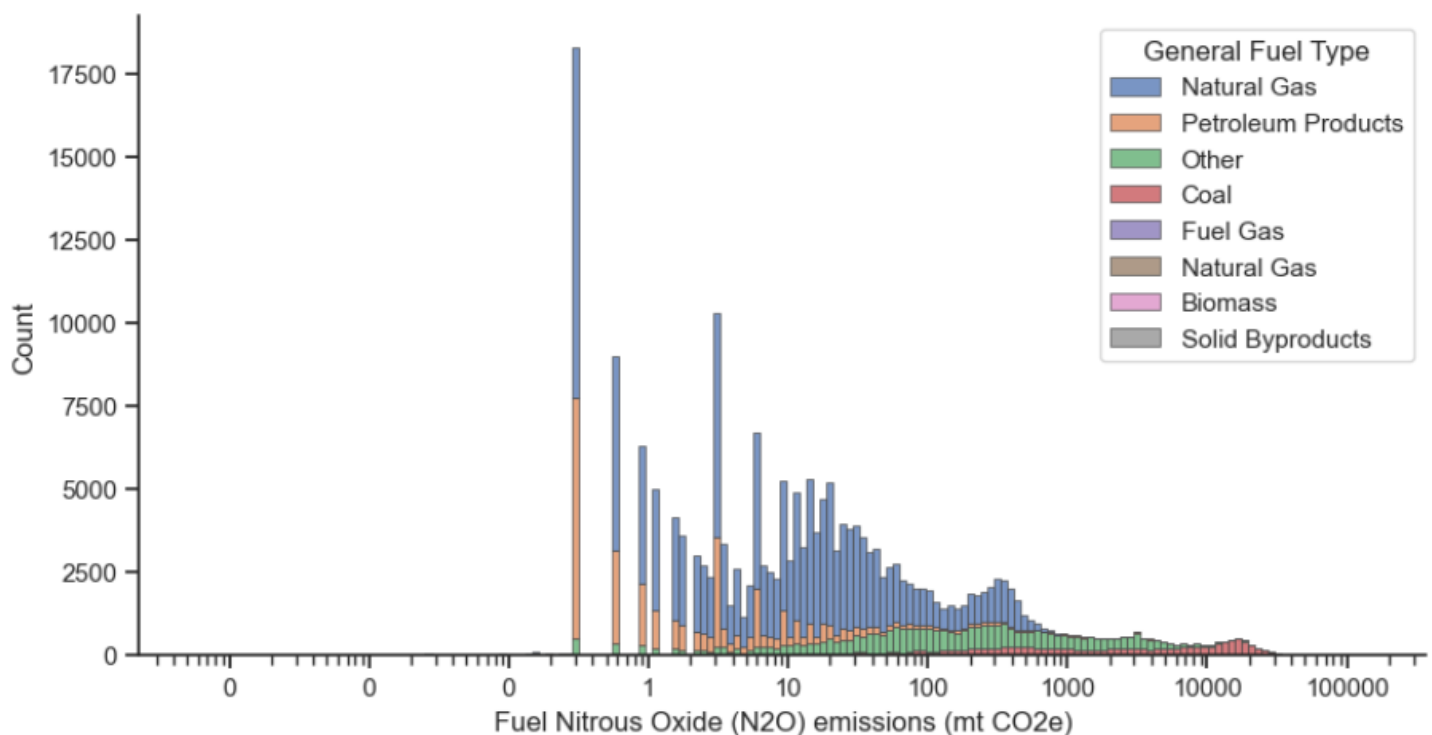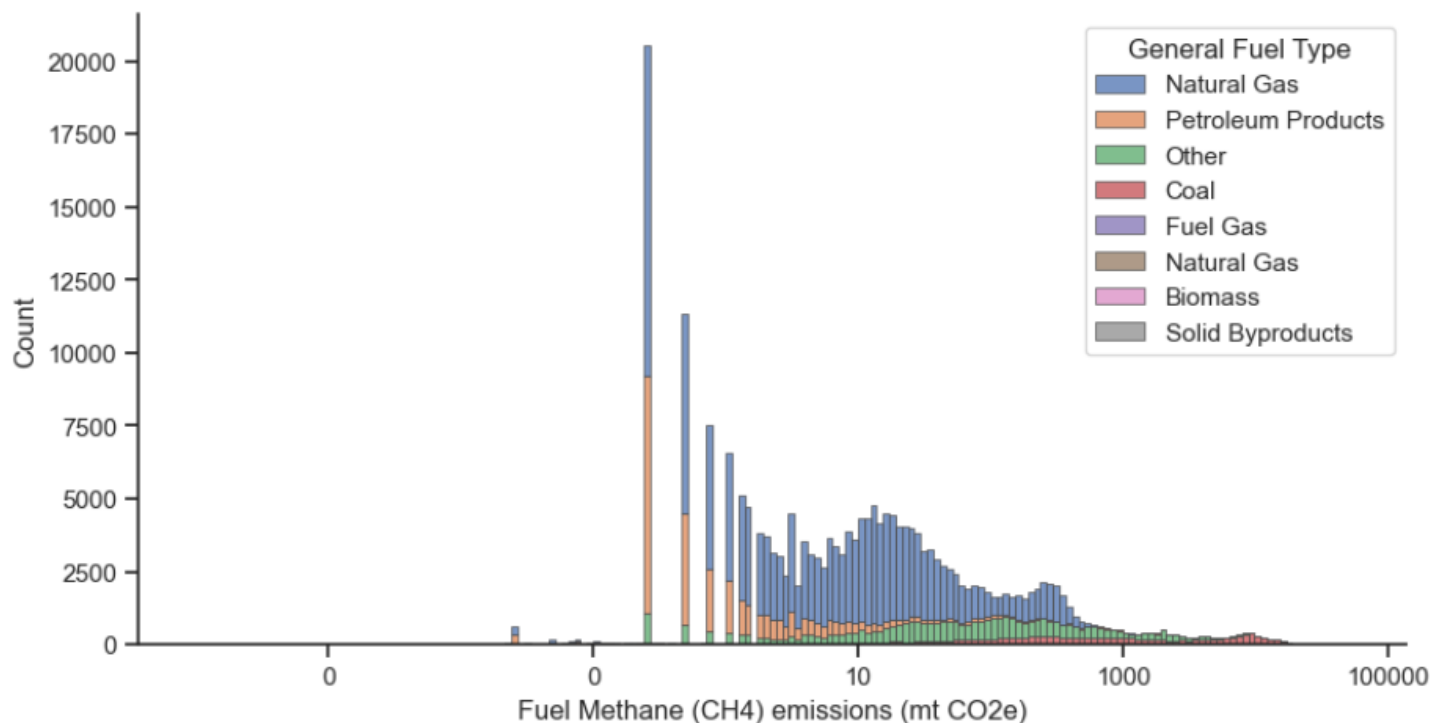
With only a slight disparity between fuel nitrous oxide (N2O) and fuel methane (CH4) emissions by general fuel type. Visually representing the dominant number of natural gas utilization. But with minimal count of coal has both achieved the maximum values for both fuel emissions.

Exploratory Data Analysis for Fuel Emissions

Proponent: Engr. Matthew David M. Loquinerio, REE
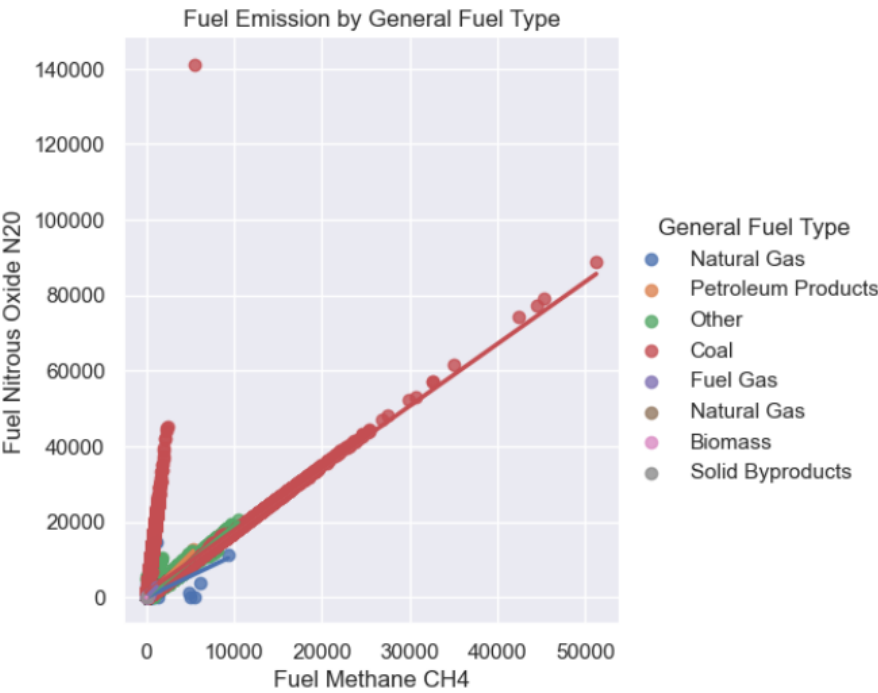
A vivid representation of fuel emission among general fuel types. Coal as a fuel type has contributed the most in emission both CH4 and N2O. While most fuel types only achieve between 10,000 CH4 and 20,000 N2O. Coal has attained values of 50,000 to nearly as 100,000 in emissions. Garnering a 500% increase ratio among general fuel types.
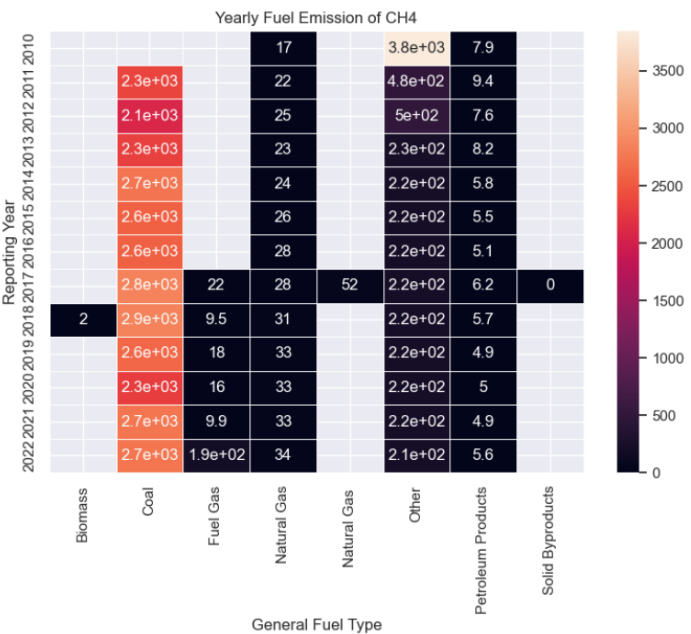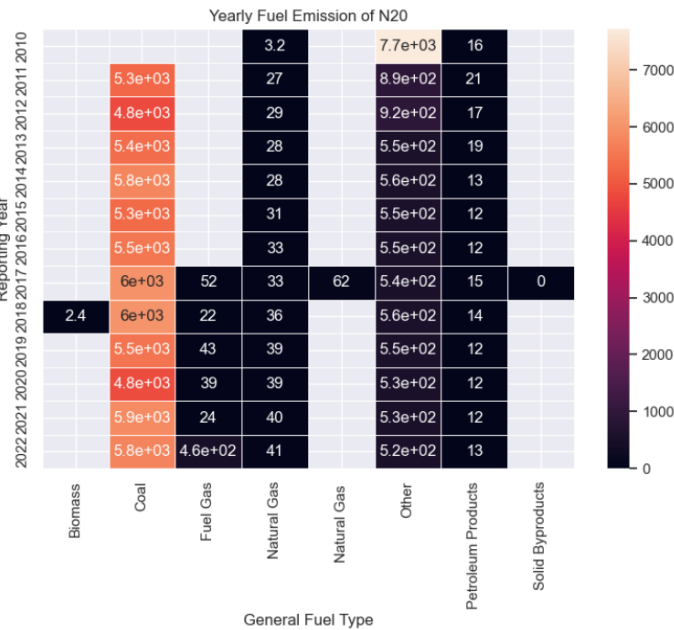
Heatmap is an ideal chart visualization to use to graphically visualize the distribution of fuel emission by general fuel type from 2010 to year 2022. Here we can infer that natural gas is at constant low-level emission. While, coal facilities that only started at year 2011 has already at high emission rate until year 2022. Also, fuel gas has a sudden uptick on value in the last year transforming a former record of 24 N2O and 9.9 CH4 to sudden above 100 emission.

Exploratory Data Analysis for Fuel Emissions

Proponent: Engr. Matthew David M. Loquinerio, REE

Sample Dashboard for Fuel Emission (2010 – 2022)