GOVERNMENT COLLEGE OF TECHNOLOGY, COIMBATORE-13

DEPARTMENT OF ECE

Comprehensive Analysis of Air Quality Data in Tamil Nadu

PHASE-V

TEAM MEMBERS

71772114140-Sindhuja R

```
# Importing the necessary libraries

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt


# Load and preprocess the air quality dataset

data = pd.read_csv('your_air_quality_data.csv')  # Replace 'your_air_quality_data.csv' with the actual dataset path


# Display a brief overview of the dataset

print("Description of columns:\n")

print(data.describe())  # Description of each column


# Check for missing data

print("\nNo. of Null Columns:\n", data.isnull().sum())  # Count of null values in columns


# Remove rows with missing data

data = data.dropna()  # To remove rows with null values


# Data preprocessing and analysis can continue from here, including data visualization, statistical analysis, and more.
```

Description of columns:

	column_1	column_2	column_3	...	column_n
count	N	N	N	...	N
mean	Mean	Mean	Mean	...	Mean
std	Std	Std	Std	...	Std
min	Min	Min	Min	...	Min
25%	25th %ile	25th %ile	25th %ile	...	25th %ile
50%	Median	Median	Median	...	Median
75%	75th %ile	75th %ile	75th %ile	...	75th %ile

max Max Max Max ... Max

No. of Null Columns:

column_1 0

column_2 5

column_3 0

...

column_n 10

dtype: int64The "Description of columns" section provides statistics such as count, mean, standard deviation, minimum, maximum, and quartiles for each numeric column in your dataset.The "No. of Null Columns" section shows the number of null (missing) values in each column.After running this code, the dataset will be preprocessed to remove rows with missing data, and you can continue with your air quality analysis and data visualization as needed.