

Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-packages (1.6.1)

Requirement already satisfied: matplotlib in /usr/local/lib/python3.11/dist-packages (3.10.0)

Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (2.2.2)

Requirement already satisfied: seaborn in /usr/local/lib/python3.11/dist-packages (0.13.2)

Requirement already satisfied: numpy>=1.19.5 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (2.0.2)

Requirement already satisfied: scipy>=1.6.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (1.15.2)

Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (1.4.2)

Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (3.6.0)

Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (1.3.2)

Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (0.12.1)

Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (4.57.0)

Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (1.4.8)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (24.2)

Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (11.2.1)

Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (3.2.3)

Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)

Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)

Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.7->matplotlib) (1.17.0)

No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving air quality database .csv to air quality database (2).csv

Dataset Shape: (12351, 11)

	Stn Code	Sampling Date	State	City/Town/Village/Area \
0	38	5/1/2010	Tamil Nadu	Chennai
1	38	7/1/2010	Tamil Nadu	Chennai
2	38	12/1/2010	Tamil Nadu	Chennai
3	38	1/19/2010	Tamil Nadu	Chennai
4	38	1/21/2010	Tamil Nadu	Chennai

Location of Monitoring Station \

0 Kathivakkam, Municipal Kalyana Mandapam, Chennai
 1 Kathivakkam, Municipal Kalyana Mandapam, Chennai

- 2 Kathivakkam, Municipal Kalyana Mandapam, Chennai
- 3 Kathivakkam, Municipal Kalyana Mandapam, Chennai
- 4 Kathivakkam, Municipal Kalyana Mandapam, Chennai

	Agency	Type of Location	SO2	NO2 \
0	Tamilnadu State Pollution Control Board	Industrial Area	9.60	17.166667
1	Tamilnadu State Pollution Control Board	Industrial Area	11.15	20.283333
2	Tamilnadu State Pollution Control Board	Industrial Area	12.45	20.516667
3	Tamilnadu State Pollution Control Board	Industrial Area	10.75	18.183333
4	Tamilnadu State Pollution Control Board	Industrial Area	9.78	17.320000

	RSPM/PM10	SPM
0	73.333333	149.666667
1	61.333333	150.333333
2	75.000000	114.666667
3	120.000000	197.666667
4	96.500000	216.000000

	Stn Code	SO2	NO2	RSPM/PM10	SPM
count	12351.000000	12149.000000	12153.000000	12303.000000	1902.000000
mean	383.105821	10.820691	21.409154	68.310122	116.251122
std	259.556466	9.897454	10.200971	45.745925	93.565013
min	38.000000	1.000000	0.000000	3.000000	4.000000
25%	237.000000	7.000000	16.000000	41.000000	54.000000
50%	307.000000	10.133333	21.000000	56.000000	98.166667
75%	761.000000	14.000000	25.000000	85.000000	151.000000
max	773.000000	909.000000	315.000000	1183.500000	1402.500000

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 12351 entries, 0 to 12350

Data columns (total 11 columns):

#	Column	Non-Null Count	Dtype
0	Stn Code	12351 non-null	int64
1	Sampling Date	12351 non-null	object
2	State	12351 non-null	object
3	City/Town/Village/Area	12351 non-null	object
4	Location of Monitoring Station	12351 non-null	object
5	Agency	12351 non-null	object
6	Type of Location	12351 non-null	object
7	SO2	12149 non-null	float64
8	NO2	12153 non-null	float64
9	RSPM/PM10	12303 non-null	float64
10	SPM	1902 non-null	float64

dtypes: float64(4), int64(1), object(6)

memory usage: 1.0+ MB

None

Index(['Stn Code', 'Sampling Date', 'State', 'City/Town/Village/Area',
'Location of Monitoring Station', 'Agency', 'Type of Location', 'SO2',
'NO2', 'RSPM/PM10', 'SPM'],

dtype='object')