

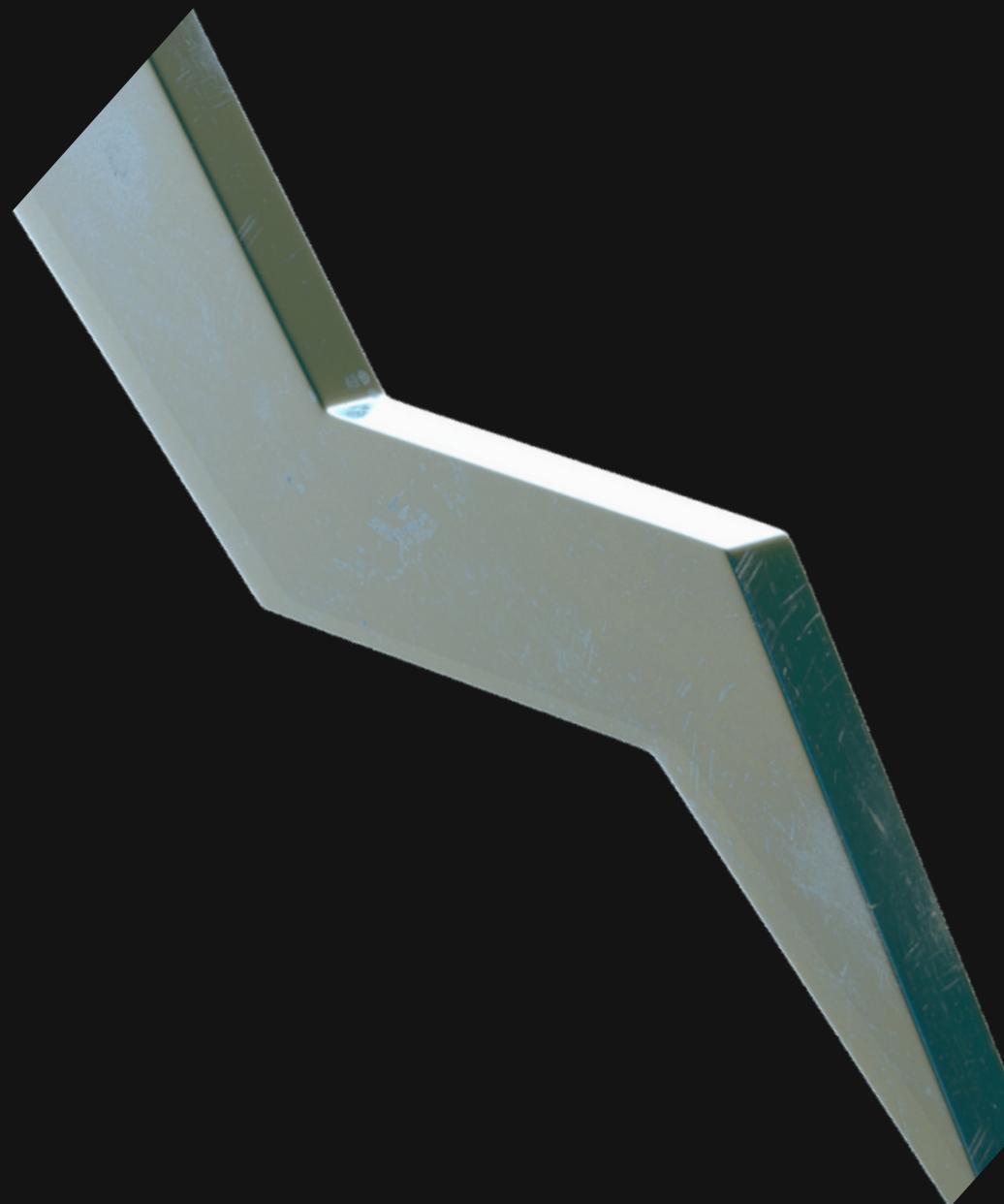
Air Quality Analysis in Tamil Nadu

TEAM MEMBERS

Sandhiya A
Priyadarshini.

B

Logaharshini.S
Kirthana.M



Untitled2.ipynb - Colaboratory ✘ WhatsApp ✘ Location wise daily Ambient Air ✘ + colab.research.google.com/drive/1MVOMI3Vry5svgmbadC7_uZtbOG-GxO2O#scrollTo=4pnOGon69UKD

Comment Share

File Edit View Insert Runtime Tools Help All changes saved

Files

+ Code + Text

```
[16] import pandas as pd
import numpy as np
import seaborn as sns
import plotly.express as px
import matplotlib.pyplot as plt

#skLearn
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split

from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
```

```
[2] df = pd.read_csv("/content/cpcb_dly_aq_tamil_nadu-2014.csv")
df
```

	Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10	PM 2.5
0	38	01-02-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.0	17.0	55.0	NaN
1	38	01-07-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	17.0	45.0	NaN
2	38	21-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.0	18.0	50.0	NaN
3	38	23-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.0	16.0	46.0	NaN
4	38	28-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	14.0	42.0	NaN
...
2874	773	12-03-14	Tamil Nadu	Trichy	Central Bus Stand, Trichy	Tamilnadu State Pollution Control Board	Residential, Rural and other Areas	15.0	18.0	102.0	NaN
2875	773	12-10-14	Tamil Nadu	Trichy	Central Bus Stand, Trichy	Tamilnadu State Pollution Control Board	Residential, Rural and other Areas	12.0	14.0	91.0	NaN
2876	773	17-12-14	Tamil Nadu	Trichy	Central Bus Stand, Trichy	Tamilnadu State Pollution Control Board	Residential, Rural and other Areas	19.0	22.0	100.0	NaN
2877	773	24-12-14	Tamil Nadu	Trichy	Central Bus Stand, Trichy	Tamilnadu State Pollution Control Board	Residential, Rural and other Areas	15.0	17.0	95.0	NaN
2878	773	31-12-14	Tamil Nadu	Trichy	Central Bus Stand, Trichy	Tamilnadu State Pollution Control Board	Residential, Rural and other Areas	14.0	16.0	94.0	NaN

2879 rows × 11 columns

```
[8] df.columns
```

```
Index(['Stn Code', 'Sampling Date', 'State', 'City/Town/Village/Area', 'Location of Monitoring Station', 'Agency', 'Type of Location', 'SO2', 'NO2', 'RSPM/PM10', 'PM 2.5'], dtype='object')
```

Disk 80.88 GB available

Os completed at 2:17PM

Untitled2.ipynb - Colaboratory ✘ WhatsApp ✘ Location wise daily Ambient Air ✘ +

← → C 🔍 colab.research.google.com/drive/1MVOMI3Vry5svgmbadC7_uZtbOG-GxO2O#scrollTo=4pnOGon69UKD

Comment Share ⚙

Untitled2.ipynb

File Edit View Insert Runtime Tools Help Saving...

Files

sample_data cpcb_dly_aq_tamil_n...

+ Code + Text

[8] Index(['Stn Code', 'Sampling Date', 'State', 'City/Town/Village/Area', 'Location of Monitoring Station', 'Agency', 'Type of Location', 'SO2', 'NO2', 'RSPM/PM10', 'PM 2.5'],
dtype='object')

[9] df.describe()

	Stn Code	SO2	NO2	RSPM/PM10	PM 2.5
count	2879.000000	2868.000000	2866.000000	2875.000000	0.0
mean	475.750261	11.503138	22.136776	62.494261	NaN
std	277.875577	5.051702	7.128694	31.368745	NaN
min	38.000000	2.000000	5.000000	12.000000	NaN
25%	238.000000	8.000000	17.000000	41.000000	NaN
50%	368.000000	12.000000	22.000000	55.000000	NaN
75%	764.000000	15.000000	25.000000	78.000000	NaN
max	773.000000	49.000000	71.000000	269.000000	NaN

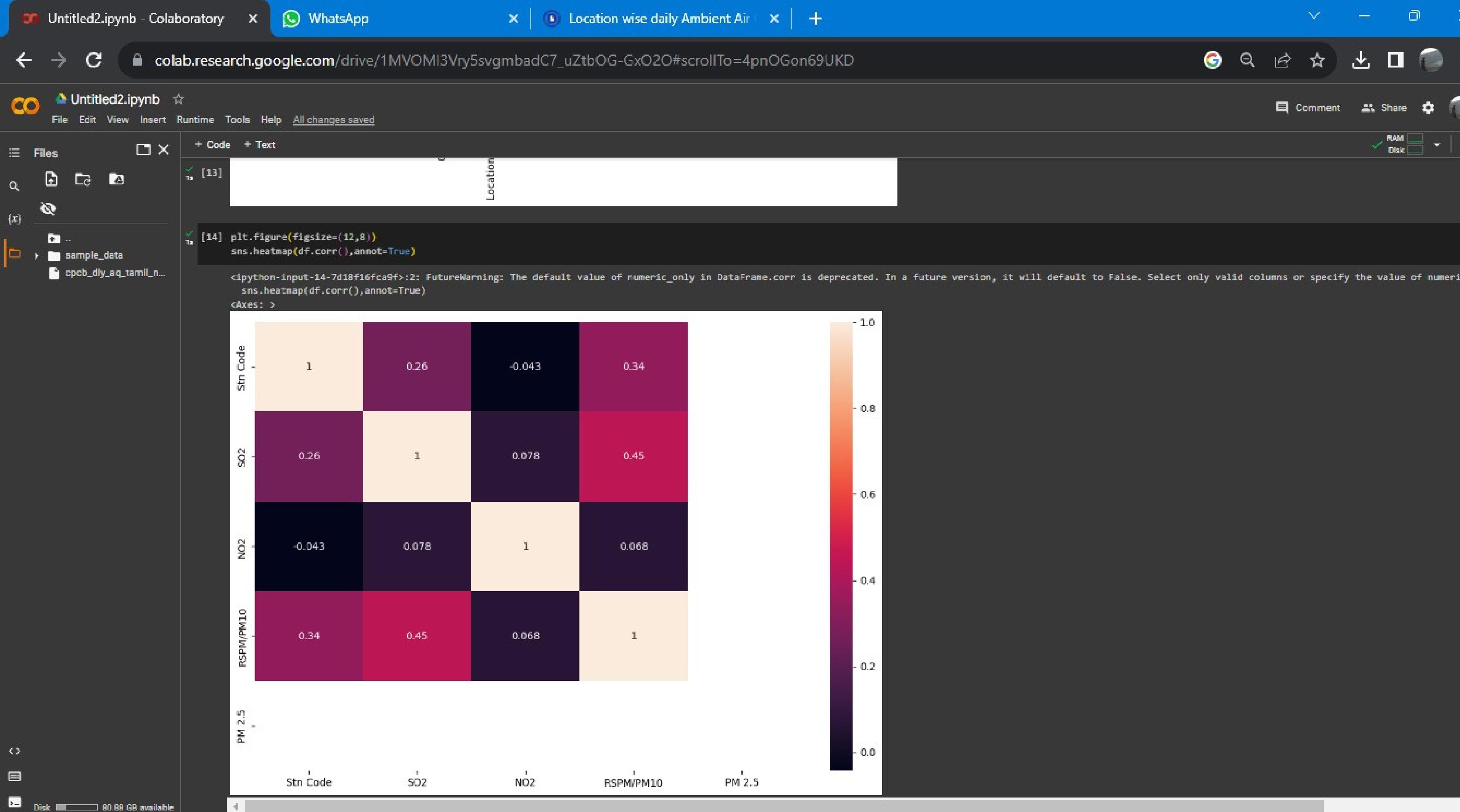
[10] df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2879 entries, 0 to 2878
Data columns (total 11 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Stn Code        2879 non-null    int64  
 1   Sampling Date   2879 non-null    object  
 2   State           2879 non-null    object  
 3   City/Town/Village/Area  2879 non-null    object  
 4   Location of Monitoring Station 2879 non-null    object  
 5   Agency          2879 non-null    object  
 6   Type of Location 2879 non-null    object  
 7   SO2             2868 non-null    float64 
 8   NO2             2866 non-null    float64 
 9   RSPM/PM10       2875 non-null    float64 
 10  PM 2.5          0 non-null      float64 
dtypes: float64(4), int64(1), object(6)
memory usage: 247.5+ KB
```

[11] df.isnull().sum()

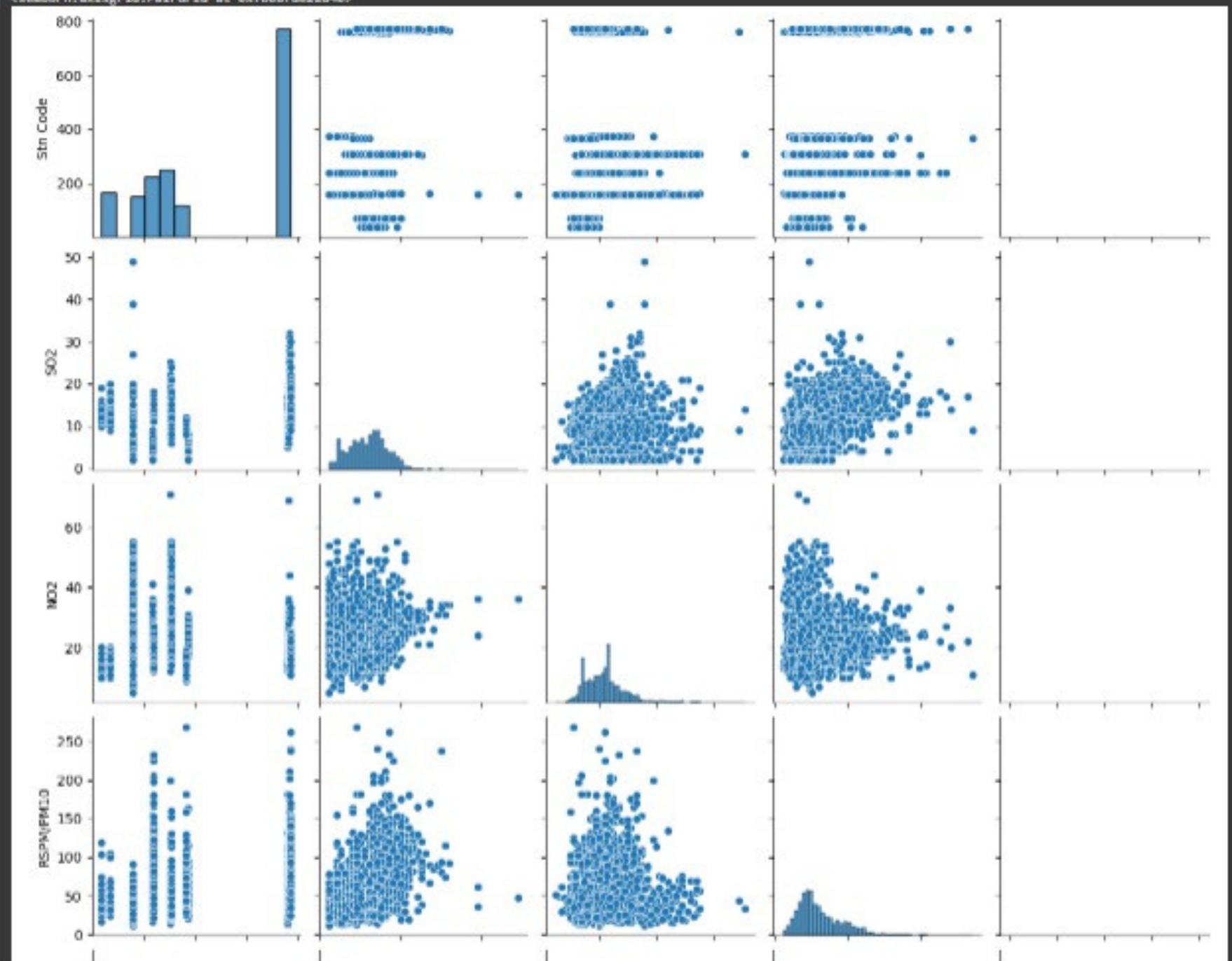
Stn Code	0
Disk	80.88 GB available

0s completed at 2:17PM

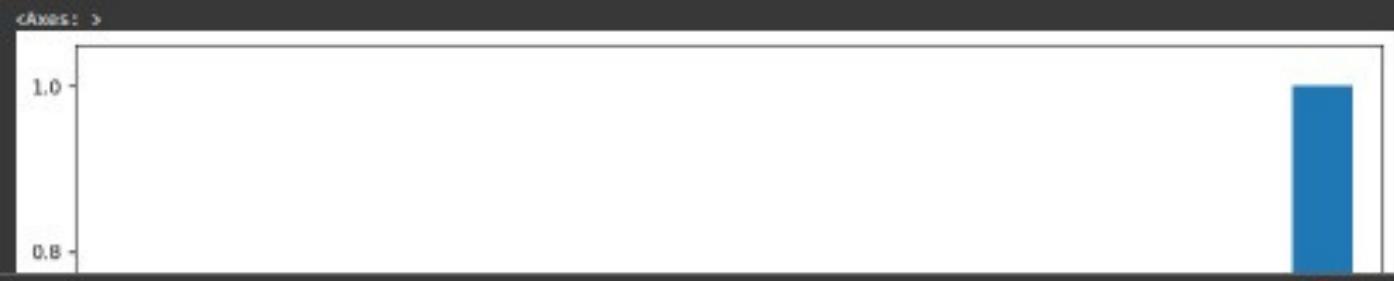


Untitled2.ipynb

Comment Share



```
[24] dF.isnull().mean().plot.bar(figsize = (12,8))
```



Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

+ Code + Text

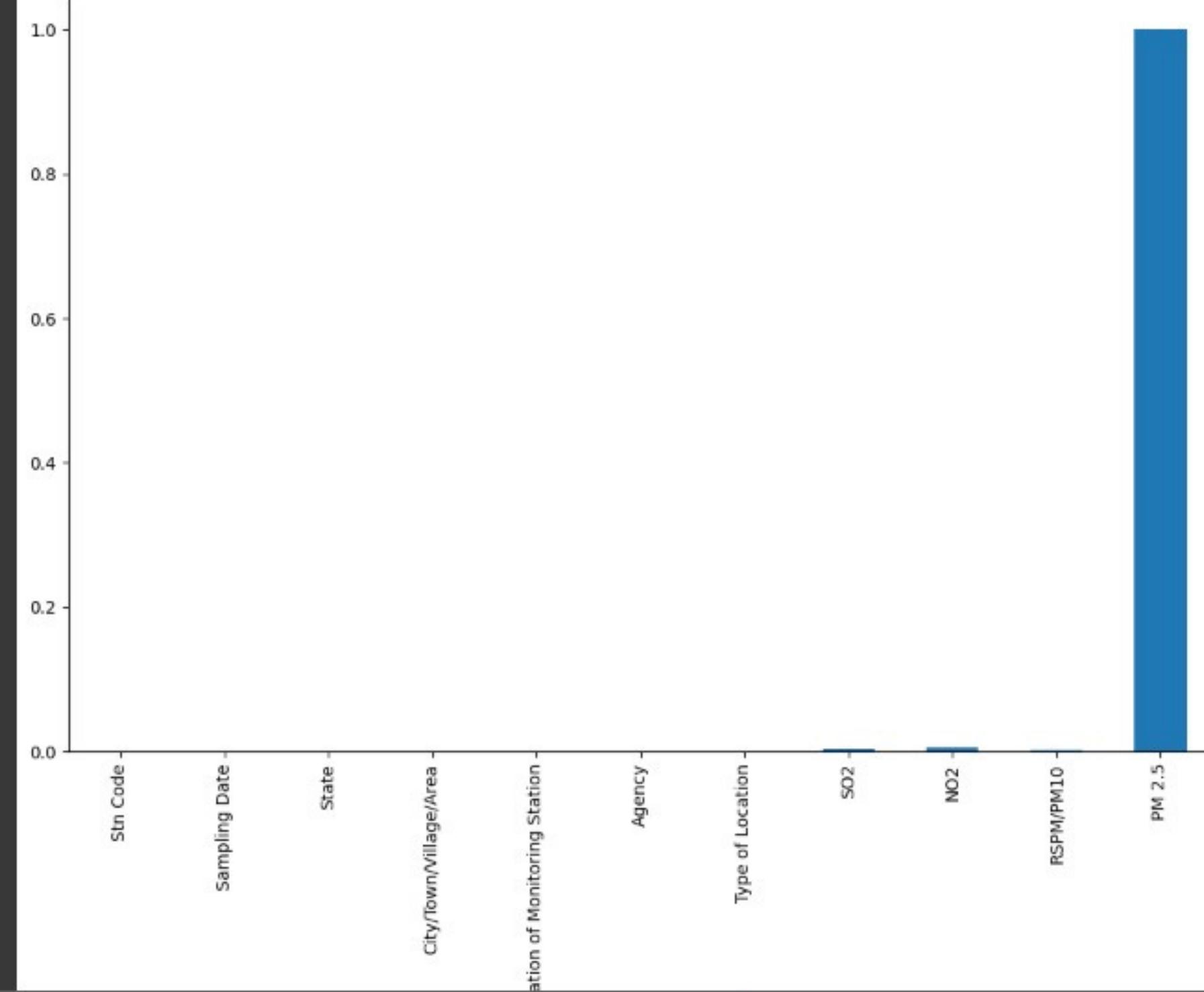
Q D C A

{x}

sample_data
cpcb_dly_aq_tamil_n...

[24] df.isnull().mean().plot.bar(figsize = (12,8))

<Axes: >



Untitled2.ipynb - Colaboratory x WhatsApp x Location wise daily Ambient Air x +

colab.research.google.com/drive/1MVOMI3Vry5svgmbadC7_uZtbOG-GxO2O#scrollTo=eGatn6dh3qhB

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data cpcb_dly_aq_tamil_n...

+ Code + Text

```
[26]: df.isnull().mean().plot.bar(figsize = (10,6))
plt.xlabel("Features")
```

Text(0.5, 8, 'Features')

0.0 0.2 0.4 0.6 0.8 1.0

Strn Code Sampling Date State City/Town/Village/Area Location of Monitoring Station Agency Type of Location SO2 NO2 RSPM/PM10 PM 2.5

Features

Disk 80.88 GB available [27]: df.isnull().sum()

0s completed at 2:17PM

Untitled2.ipynb - Colaboratory x WhatsApp x Location wise daily Ambient Air x +

colab.research.google.com/drive/1MVOMI3Vry5svgmbadC7_uZtbOG-GxO2O#scrollTo=eGatn6dh3qhB

Comment Share

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

+ Code + Text

[27] df.isnull().sum()

```
Stn Code      0
Sampling Date 0
State         0
City/Town/Village/Area 0
Location of Monitoring Station 0
Agency        0
Type of Location 0
SO2           11
NO2           13
RSPM/PM10     4
PM 2.5        2879
dtype: int64
```

[28] sns.heatmap(df.isnull())

<Axes: >

Disk 80.88 GB available

0s completed at 2:17PM

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- (x) sample_data
- cpcb_dly_aq_tamil_n...

+ Code + Text

[28]

[29] df.head()

	Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10	PM 2.5
0	38	01-02-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	11.0	17.0	55.0	NaN
1	38	01-07-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	17.0	45.0	NaN
2	38	21-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	12.0	18.0	50.0	NaN
3	38	23-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	15.0	16.0	48.0	NaN
4	38	28-01-14	Tamil Nadu	Chennai	Kathivakkam, Municipal Kalyana Mandapam, Chennai	Tamilnadu State Pollution Control Board	Industrial Area	13.0	14.0	42.0	NaN

Disk 80.88 GB available