| | GRT INSTITUTE OF ENGINEERING AND TECHNOLOGY, TIRUTTANI - 631209<br>Approved by AICTE, New Delhi Affiliated to Anna University, Chennai | |
|---|---|---|

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**Phase4**

**PROJECT TITLE**

# *Air quality analysis and prediction in tamilnadu*

**COLLEGE CODE:1103**

**Janani.L**

3rd yr, 5th sem

Reg no:110321104018

**janani26dec2003@gmail.com**

Program:

```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt

# Load the air quality data into a DataFrame
dtype = {
     'stn_code': str,
    'sampling_date': str,
    'state': str,
    'location': str,
    'agency':str,
    'type':str,
    'so2':float,
    'no2':float,
    'rspm':float,
    'spm':float,
    'location_monitoring_station':str,
    'date':str
}
data= pd.read_csv('C:/project/data.csv', sep=',', encoding='ISO-8859-1',
dtype=dtype)
data=data.sample(500)
print(data)
# Calculate average levels
average_so2 = data['so2'].mean()
average_no2 = data['no2'].mean()
average_rspm = data['rspm'].mean()

# Identify pollution trends
# You may want to add more analysis here

# Create visualizations
plt.figure(figsize=(10, 6))
plt.bar(['so2', 'no2', 'rspm'], [average_so2, average_no2, average_rspm])
plt.xlabel('Pollutant')
plt.ylabel('Average Level')
plt.title('Average Air Quality Levels')
plt.show()
# Create a bar chart to visualize the averages
pollutants = ['SO2', 'NO2', 'RSPM']
averages = [average_so2,average_no2 , average_rspm]
plt.bar(pollutants, averages)
plt.xlabel('Pollutant')
plt.ylabel('Average Level')
plt.title('Average Air Quality Levels')
plt.show()
# Calculate average levels for SO2, NO2, and RSPM/PM10
average_so2 = data['so2'].mean()
average_no2 = data['no2'].mean()
average_rspm_pm10 = data['rspm'].mean()

print(f"Average SO2 level: {average_so2}")
print(f"Average NO2 level: {average_no2}")
print(f"Average RSPM/PM10 level: {average_rspm}")
```

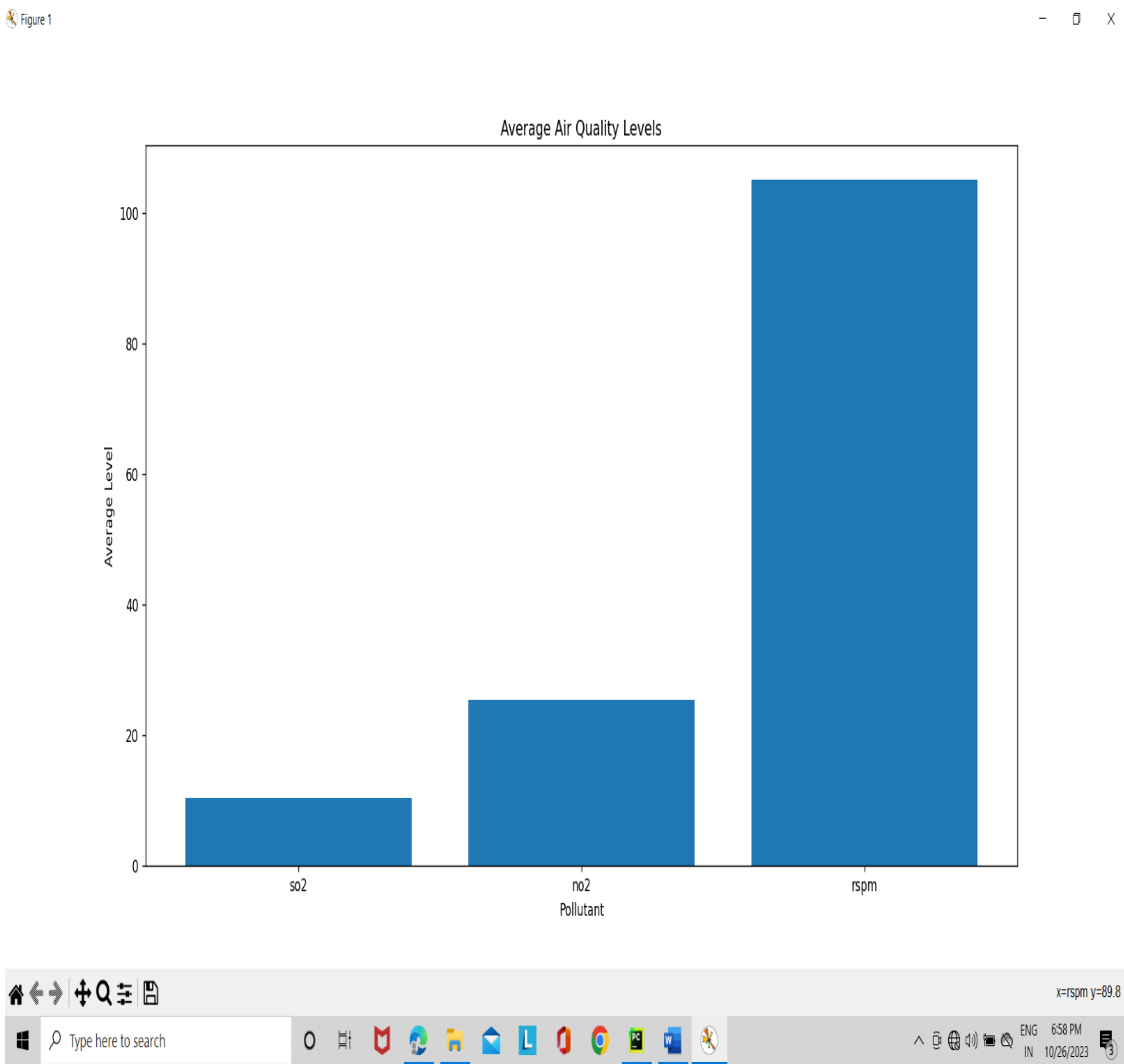
```python
plt.figure(figsize=(12, 6))
plt.plot(data['sampling_date'], data['so2'], label='SO2 Levels',
color='blue')
plt.xlabel('sampling_date')
plt.ylabel('SO2 Levels')
plt.title('SO2 Levels in Tamil Nadu')
plt.legend()
plt.grid()
plt.show()
plt.figure(figsize=(12, 6))
plt.plot(data['sampling_date'], data['no2'], label='no2 Levels',
color='violet')
plt.xlabel('sampling_date')
plt.ylabel('no2 Levels')
plt.title('no2 Levels in Tamil Nadu')
plt.legend()
plt.grid()
plt.show()
plt.figure(figsize=(12, 6))
plt.plot(data['sampling_date'], data['rspm'], label='rspm Levels',
color='black')
plt.xlabel('sampling_date')
plt.ylabel('rspm Levels')
plt.title('rspm Levels in Tamil Nadu')
plt.legend()
plt.grid()
plt.show()

import seaborn as sns
# Basic histograms for SO2, NO2, and RSPM/PM10
plt.figure(figsize=(12, 6))
plt.subplot(131)
sns.histplot(data['so2'], bins=20, kde=True)
plt.title('SO2 Distribution')

plt.subplot(132)
sns.histplot(data['no2'], bins=20, kde=True)
plt.title('NO2 Distribution')

plt.subplot(133)
sns.histplot(data['rspm'], bins=20, kde=True)
plt.title('RSPM/PM10 Distribution')
plt.tight_layout()
plt.show()

# Calculate average levels for SO2, NO2, and RSPM/PM10
average_so2 = data.groupby('location_monitoring_station')['so2'].mean()
average_no2 = data.groupby('location_monitoring_station')['no2'].mean()
average_rspm_pm10 =
data.groupby('location_monitoring_station')['rspm'].mean()

# Create visualizations
# Example: Bar chart to visualize average SO2 levels by city
plt.figure(figsize=(10, 6))
plt.bar(average_so2.index, average_so2.values)
plt.xlabel('City')
plt.ylabel('Average SO2 Level')
```

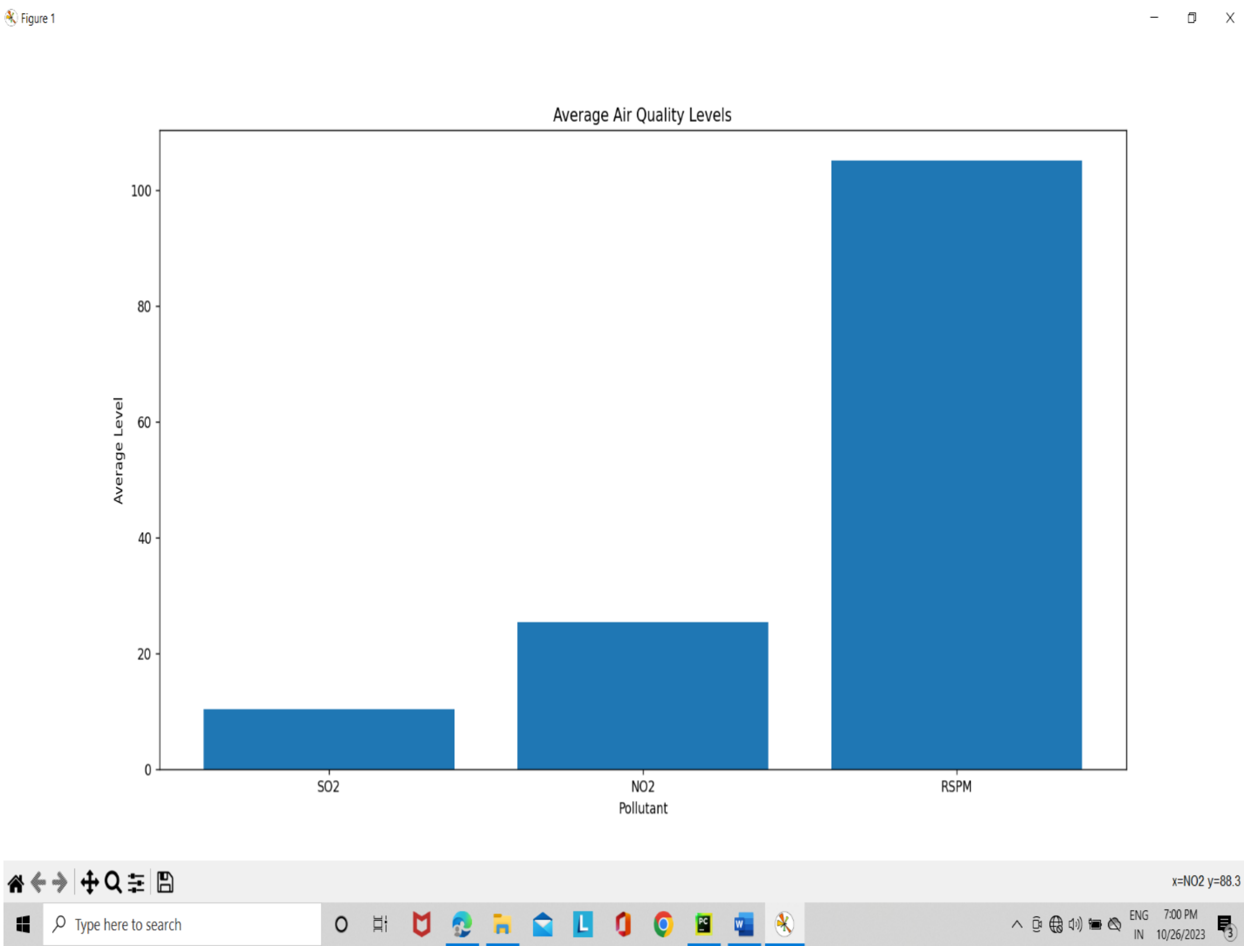
```
plt.title('Average SO2 Levels by City')
plt.xticks(rotation=45)
plt.show()

# Calculate average SO2, NO2, and RSPM/PM10 levels
avg_so2 = data.groupby('location_monitoring_station')['so2'].mean()
avg_no2 = data.groupby('location_monitoring_station')['no2'].mean()
avg_rspm_pm10 = data.groupby('location_monitoring_station')['rspm'].mean()

# Identify areas with high pollution levels
high_so2_areas = avg_so2[avg_so2 > 2]  # Replace 'threshold' with your
desired threshold value
high_no2_areas = avg_no2[avg_no2 > 200]
high_rspm_pm10_areas = avg_rspm_pm10[avg_rspm_pm10 >60 ]
# Print or visualize the results
print("Average SO2 levels by location:")
print(avg_so2)
print("\nAverage NO2 levels by location:")
print(avg_no2)
print("\nAverage RSPM/PM10 levels by location:")
print(avg_rspm_pm10)

print("\nAreas with high SO2 levels:")
print(high_so2_areas)
print("\nAreas with high NO2 levels:")
print(high_no2_areas)
print("\nAreas with high RSPM/PM10 levels:")
print(high_rspm_pm10_areas)

# Visualize pollution trends over time
import matplotlib.pyplot as plt
plt.figure(figsize=(12, 6))
plt.plot(high_so2_areas.index, high_so2_areas.values, label='SO2')
plt.plot(high_no2_areas.index, high_no2_areas.values, label='NO2')
plt.plot(high_rspm_pm10_areas.index,high_rspm_pm10_areas.values,
label='RSPM/PM10')
plt.xlabel('Time')
plt.ylabel('Average Levels')
plt.title('Air Quality Trends Over Time')
plt.legend()
plt.show()
```

Output:

| | stn_code | sampling_date | ... | Unnamed: 11 | date |
|---|---|---|---|---|---|
| 10031 | NaN | 17/4/2009 | ... | NaN | 4/17/2009 |
| 227291 | 165 | 18/03/2011 | ... | NaN | 3/18/2011 |
| 367294 | 581 | 1/1/2015 | ... | NaN | 1/1/2015 |
| 21563 | 742 | 4/9/2014 | ... | NaN | 9/4/2014 |
| 72170 | 144 | 20/10/2011 | ... | NaN | 10/20/2011 |

```
...      ...       ...  ...      ...      ...
351040     NaN     22/5/2009  ...       NaN   5/22/2009
395146       6    17/02/2012  ...       NaN   2/17/2012
275222     NaN     7/12/2007  ...       NaN   12/7/2007
398317       1      14-03-13  ...       NaN   3/14/2013
107544     NaN      27-10-05  ...       NaN  10/27/2005

[500 rows x 13 columns]
```






Average SO2 level: 10.347529112074236

Average NO2 level: 25.39095238093789

Average RSPM/PM10 level: 105.0656695780131

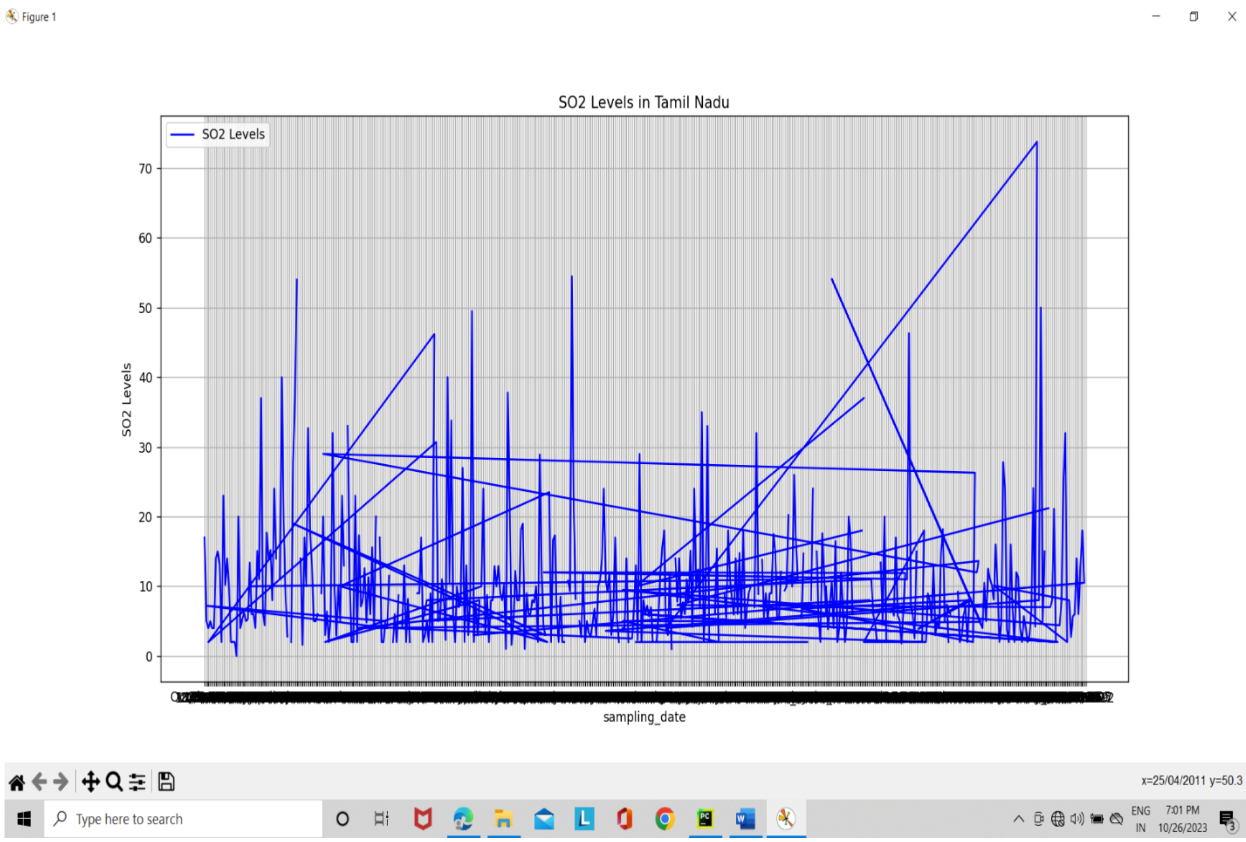


Figure 1
SO2 Levels in Tamil Nadu
SO2 Levels
70
60
50
40
30
20
10
0
SO2 Levels
sampling_date
x=25/04/2011 y=50.3
Type here to search
ENG
IN
7:01 PM
10/26/2023

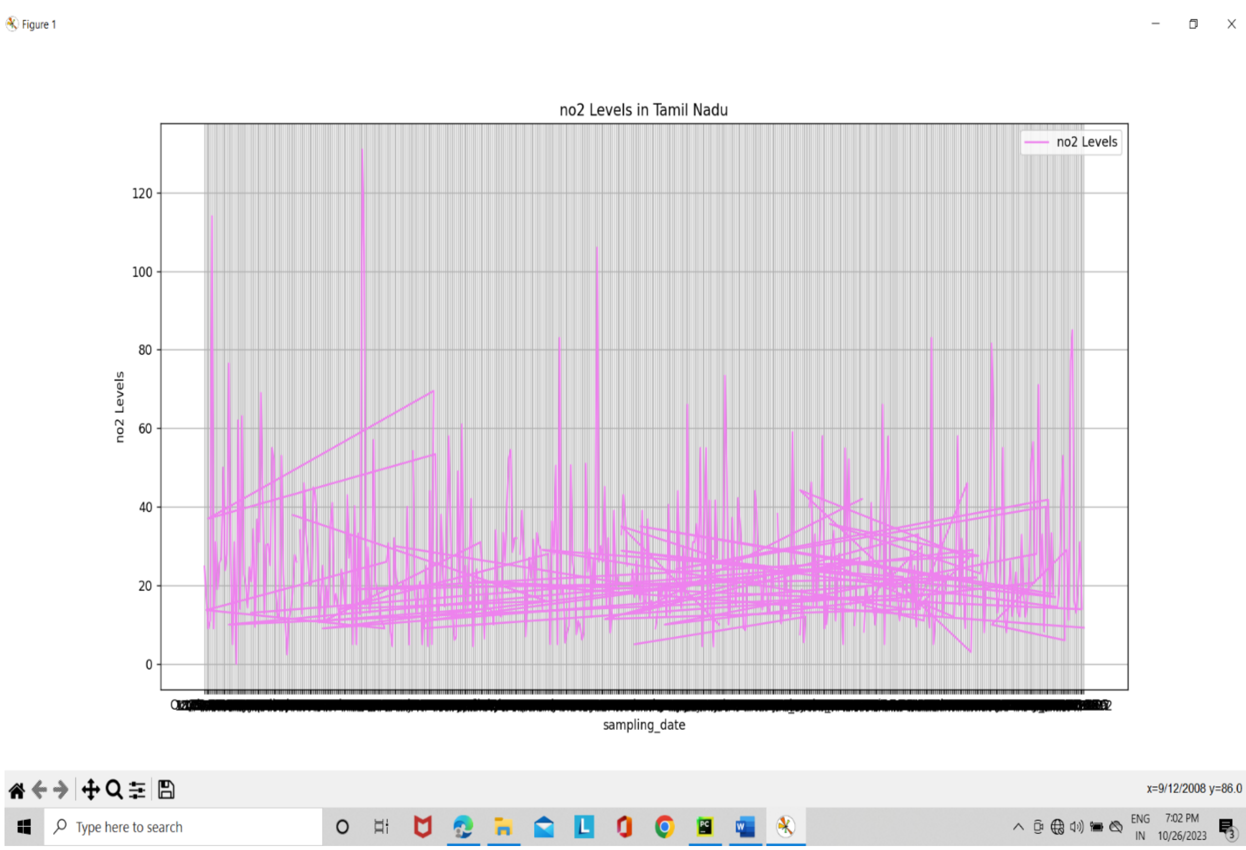


Figure 1
no2 Levels in Tamil Nadu
no2 Levels
120
100
80
60
40
20
0
no2 Levels
sampling_date
x=9/12/2008 y=86.0
Type here to search
ENG
IN
7:02 PM
10/26/2023

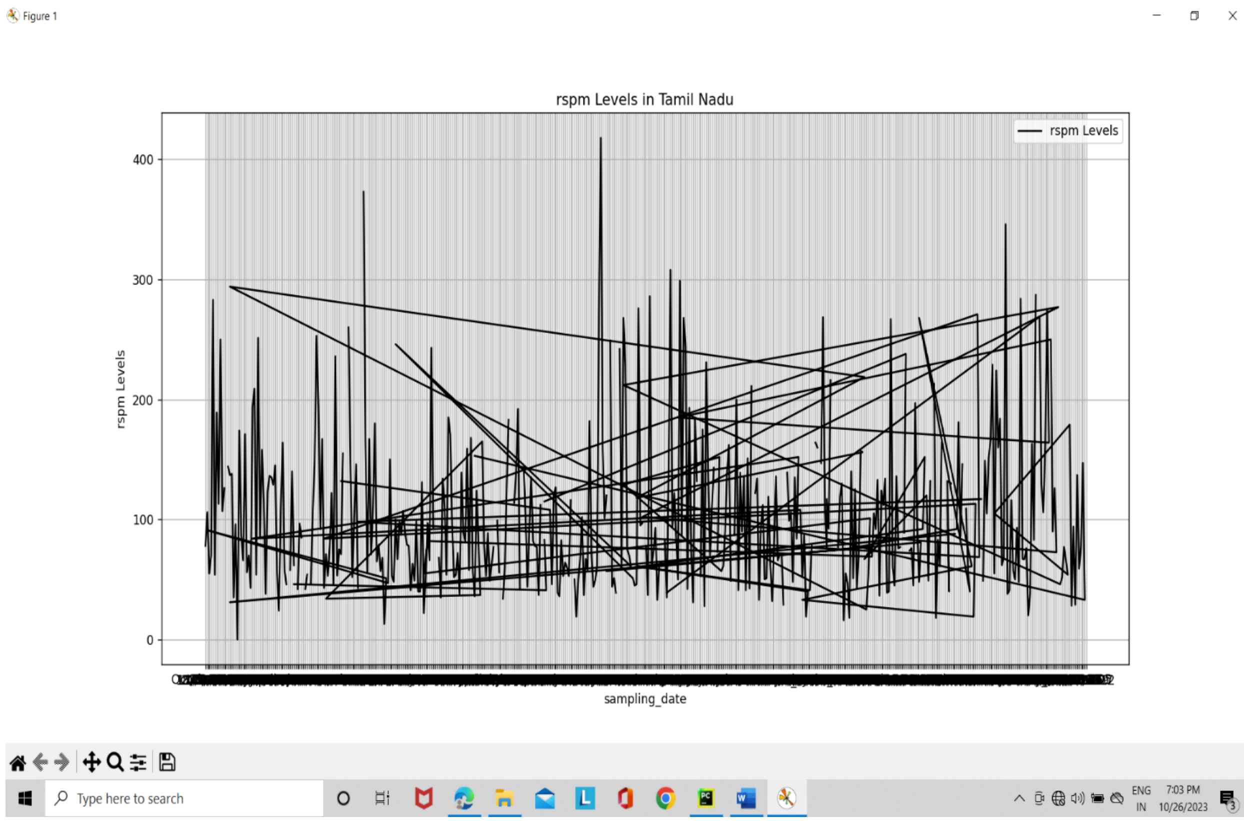


Figure 1
rspm Levels in Tamil Nadu
rspm Levels
400
300
200
100
0
rspm Levels
sampling_date
Type here to search
ENG
IN
7:03 PM
10/26/2023

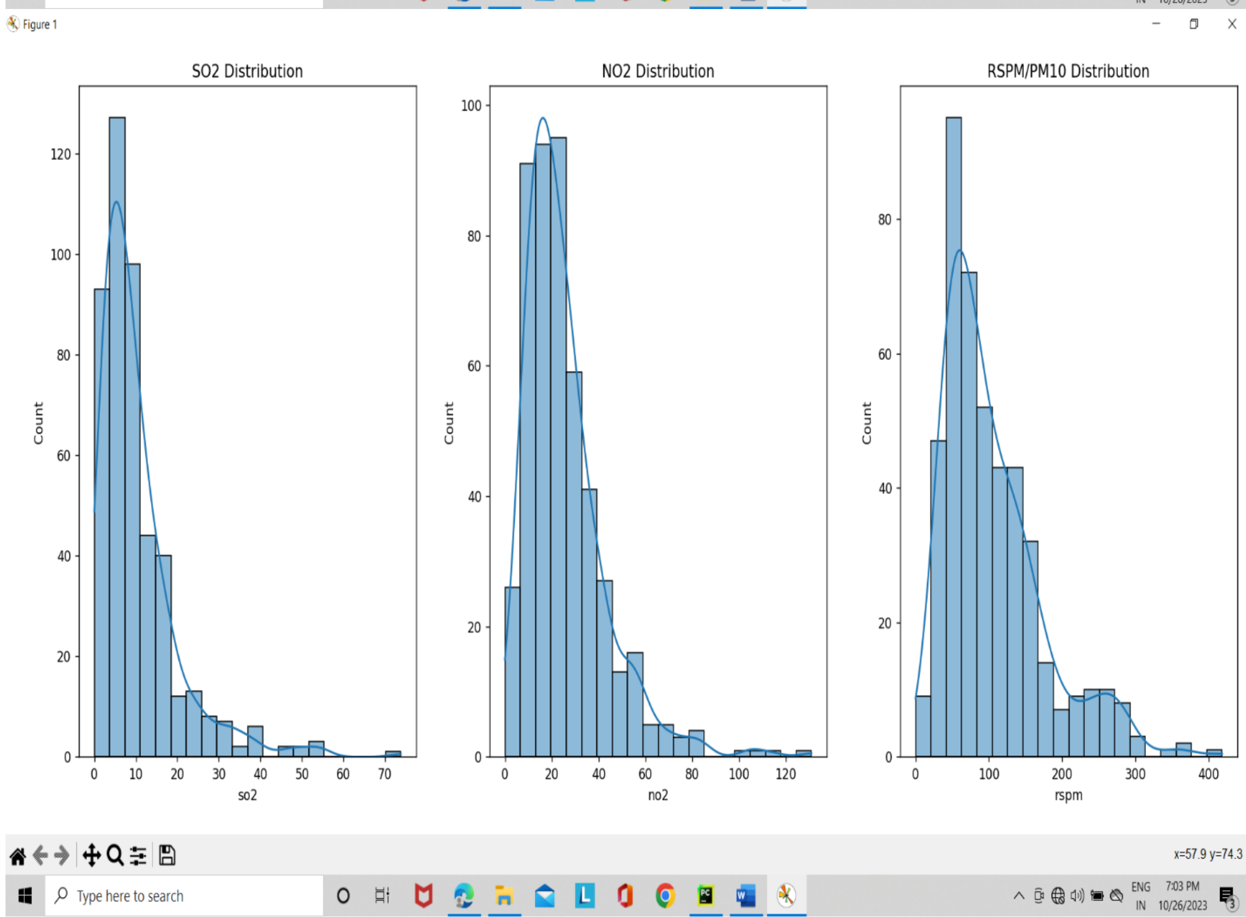
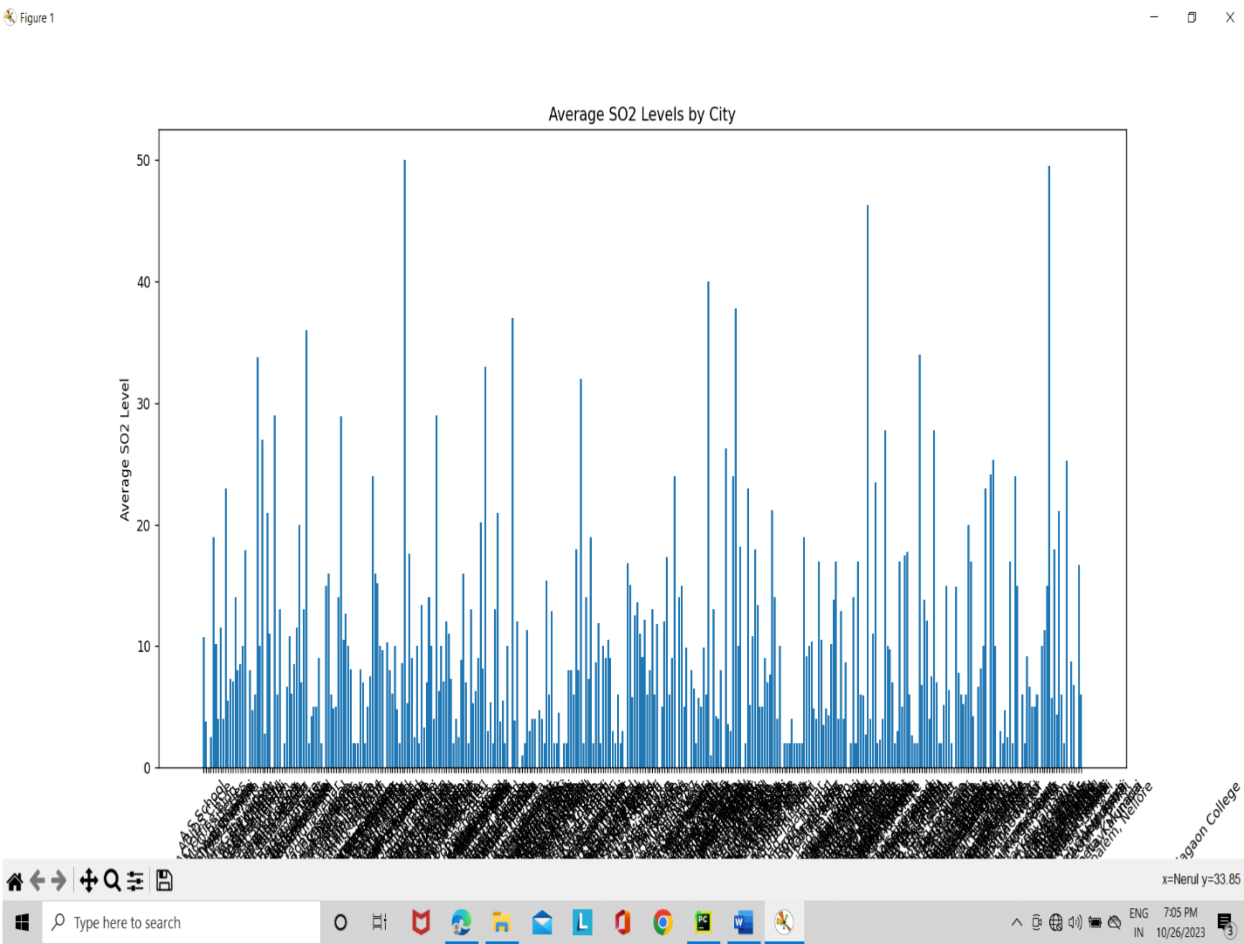


Figure 1
SO2 Distribution
NO2 Distribution
RSPM/PM10 Distribution
Count
so2
no2
rspm
0
10
20
30
40
50
60
70
80
100
120
200
300
400
x=57.9 y=74.3
Type here to search
ENG
IN
7:03 PM
10/26/2023




Average SO2 levels by location:

location_monitoring_station

| | |
|---|---|
| A S School | 10.75 |
| ABIDS Circle | 3.80 |
| AC Office Bldg. | NaN |
| AC Office Building, Parwanoo | 2.50 |
| AMCO Batteries | 19.00 |
| ... | |
| Visak Hostel, Sector-4, Bhilai | 8.75 |
| WBIIDC | 6.80 |
| WBIIDC, Haldia | NaN |
| WIT Campus | 16.70 |

```
Water Resources Division Office Campus, Christian Patty, near Nagaon College     6.00
Name: so2, Length: 359, dtype: float64

Average NO2 levels by location:
location_monitoring_station
A S School                                                  30.35
ABIDS Circle                                                 54.10
AC Office Bldg.                                               7.50
AC Office Building, Parwanoo                                          13.34
AMCO Batteries                                                  37.90
                                                         ...
Visak Hostel, Sector-4, Bhilai                                      21.25
WBIIDC                                                      44.20
WBIIDC, Haldia                                                 29.60
WIT Campus                                                     36.30
Water Resources Division Office Campus, Christian Patty, near Nagaon College    21.00
Name: no2, Length: 359, dtype: float64

Average RSPM/PM10 levels by location:
location_monitoring_station
A S School                                                  240.50
ABIDS Circle                                                  49.00
AC Office Bldg.                                                50.00
AC Office Building, Parwanoo                                          71.00
AMCO Batteries                                                  46.00
                                                         ...
Visak Hostel, Sector-4, Bhilai                                      88.64
WBIIDC                                                      33.00
WBIIDC, Haldia                                                 73.40
```

WIT Campus 121.00

Water Resources Division Office Campus, Christian Patty, near Nagaon College 194.00

Name: rspm, Length: 359, dtype: float64

Areas with high SO2 levels:

location_monitoring_station

A S School 10.75

ABIDS Circle 3.80

AC Office Building, Parwanoo 2.50

AMCO Batteries 19.00

AS School, Khanna 10.20

...

Victoria Hospital 25.30

Visak Hostel, Sector-4, Bhilai 8.75

WBIIDC 6.80

WIT Campus 16.70

Water Resources Division Office Campus, Christian Patty, near Nagaon College 6.00

Name: so2, Length: 285, dtype: float64

Areas with high NO2 levels:

Series([], Name: no2, dtype: float64)

Areas with high RSPM/PM10 levels:

location_monitoring_station

A S School 240.50

AC Office Building, Parwanoo 71.00

AS School, Khanna 229.00

ASRAM Diagnostic Center, Eluru 88.00

AZL Behrampura, Ahmadabad 66.50

...

Vijay Nagar 157.00

Visak Hostel, Sector-4, Bhilai 88.64

WBIIDC, Haldia 73.40

WIT Campus 121.00

Water Resources Division Office Campus, Christian Patty, near Nagaon College 194.00

Name: rspm, Length: 247, dtype: float64