AIR QUALITY ASSESSMENT OF TAMIL NADU

PHASE -4

- Perform the air quality analysis and create visualizations.
- Calculate average SO2, NO2, and RSPM/PM10 levels across different monitoring stations,cities, or areas. Identify pollution trends and areas with high pollution levels.
- Create visualizations using data visualization libraries (e.g., Matplotlib, Seaborn).

Calculate average SO2, NO2, and RSPM/PM10 levels

Python code:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming you have the data in a CSV file named 'pollution_data.csv'
data = pd.read_csv('project.csv')

# Calculating average SO2, NO2, and RSPM/PM10 levels
avg_SO2 = data['SO2'].mean()
avg_NO2 = data['NO2'].mean()
avg_RSPM_PM10 = data['RSPM/PM10'].mean()

# Printing average values
print(f"Average SO2 level: {avg_SO2}")
print(f"Average NO2 level: {avg_NO2}")
```

```
print(f"Average RSPM/PM10 level:
{avg_RSPM_PM10}")
```

output:

```
Average SO2 level: 9.23952006260519
```

```
Average NO2 level: 20.094601509218666
```

```
Average RSPM/PM10 level:
66.3474691699006
```

Calculating the overall average NO2 ,SO2,RSPM/PM10 across City/Town/Village/Area, Location of Monitoring Station, Type of Location

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming you have the data in a CSV
file named 'pollution_data.csv'
data = pd.read_csv('project.csv')

# Calculating average SO2, NO2, and
RSPM/PM10 levels
avg_SO2 = data['SO2'].mean()
avg_NO2 = data['NO2'].mean()
avg_RSPM_PM10 =
data['RSPM/PM10'].mean()

# Printing average values
print(f"Average SO2 level: {avg_SO2}")
print(f"Average NO2 level: {avg_NO2}")
print(f"Average RSPM/PM10 level:
{avg_RSPM_PM10}")

pollution_by_area =
data.groupby('City/Town/Village/Area')
['SO2', 'NO2', 'RSPM/PM10'].mean()
```

```
print(pollution_by_area)
pollution_by_mon =
data.groupby('Location of Monitoring
Station')['SO2', 'NO2',
'RSPM/PM10'].mean()
print(pollution_by_mon)
pollution_by_a = data.groupby('Type of
Location')['SO2', 'NO2',
'RSPM/PM10'].mean()
print(pollution_by_a)
```

Output:

```
City/Town/Village/Area
Chennai                      9.673512
15.461613   59.202182
Coimbatore                   4.913271
27.096106   62.718199
Cuddalore                    9.875000
16.312500   47.312500
Madurai                     10.976659
24.644290   47.701075
Mettur                       8.181818
24.509091   44.909091
Salem                        8.374051
26.175840   80.865854
Thoothukudi                 12.603426
13.794342  115.228898
Trichy                      12.500000
22.250000  126.625000
```

SO2 NO2 \
Location of Monitoring Station
Adyar, Chennai
12.857143 17.214286
Anna Nagar, Chennai
13.595238 19.190476
District Environmental Engineer Office,
Imperia... 9.875000 16.312500
Distt. Collector's Office, Coimbatore
4.880963 24.102556
Fenner (I) Ltd. Employees Assiciation
Building ... 11.686645 24.842953
Fisheries College, Tuticorin
13.453424 14.469138
Govt. High School, Manali, Chennai.
11.378317 18.986438
Highway (Project -I) Building, Madurai
10.405261 23.976961
Kathivakkam, Municipal Kalyana Mandapam,
Chennai 11.888405 17.514499
Kunnathur Chatram East Avani Mollai Street,
Mad... 10.735638 25.127482
Madras Medical College, Chennai
6.215530 12.087879
Main Guard Gate, Tirchy
12.500000 22.250000
NEERI, CSIR Campus Chennai
4.691930 9.827719
Poniarajapuram, On the top of DEL,
Coimbatore 3.732855 22.019043

Raja Agencies, Tuticorin
11.150523 12.640911
Raman Nagar, Mettur
7.206897 22.655172
SIDCO Industrial Complex, Mettur
9.269231 26.576923
SIDCO Office, Coimbatore
5.882353 34.502605
Sowdeswari College Building, Salem
8.374051 26.175840
Thiruvottiyur Municipal Office, Chennai
6.262057 9.944415
Thiruvottiyur, Chennai
12.552717 20.095833

RSPM/PM10
Location of Monitoring Station
Adyar, Chennai
44.190476
Anna Nagar, Chennai
97.571429
District Environmental Engineer Office,
Imperia... 47.312500
Distt. Collector's Office, Coimbatore
47.270123
Fenner (I) Ltd. Employees Assiciation
Building ... 43.407895
Fisheries College, Tuticorin
86.205215

Govt. High School, Manali, Chennai.
80.457516
Highway (Project -I) Building, Madurai
50.941176
Kathivakkam, Municipal Kalyana Mandapam,
Chennai 64.737219
Kunnathur Chatram East Avani Mollai Street,
Mad... 49.391844
Madras Medical College, Chennai
38.518939
Main Guard Gate, Tirchy
126.625000
NEERI, CSIR Campus Chennai
32.673684
Poniarajapuram, On the top of DEL,
Coimbatore 44.925532
Raja Agencies, Tuticorin
164.839147
Raman Nagar, Mettur
40.931034
SIDCO Industrial Complex, Mettur
49.346154
SIDCO Office, Coimbatore
94.298039
Sowdeswari College Building, Salem
80.865854
Thiruvottiyur Municipal Office, Chennai
38.290780
Thiruvottiyur, Chennai
83.710145

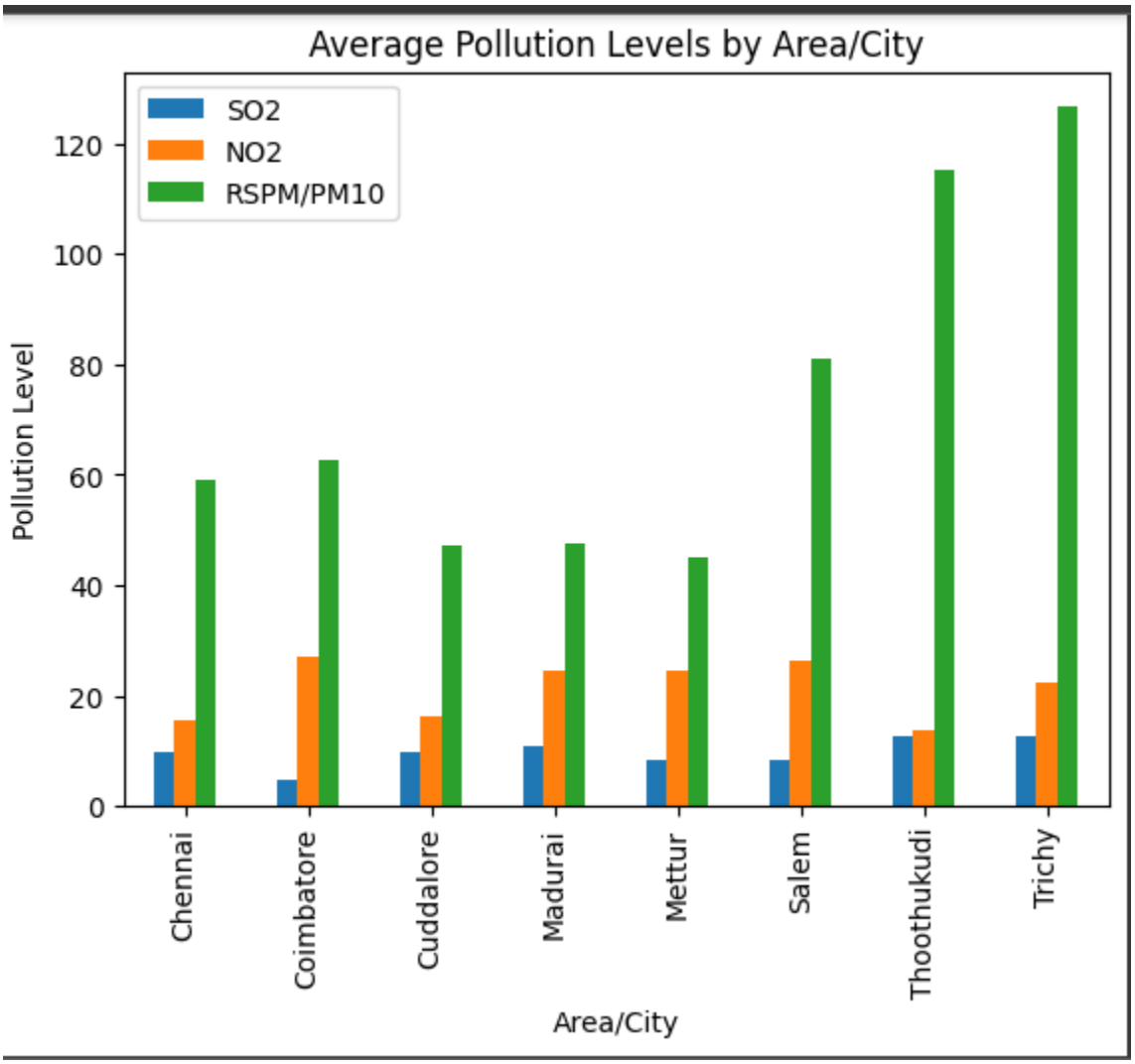
```
SO2         NO2  RSPM/PM10
Type of Location
Industrial Area
10.143347  20.466202  77.498283
Residential, Rural and other Areas
8.530706  19.803179  57.602594
```

Create visualizations using data visualization libraries SO2, NO2, and RSPM/PM10 levels across different monitoring stations, cities, or areas. Identify pollution trends and areas with high pollution levels.

Bar chart:

```
pollution_by_area =
data.groupby('City/Town/Village/Area')['SO2
', 'NO2', 'RSPM/PM10'].mean()

# Visualizing data
# Create a bar plot to show the average
pollution levels in different areas
plt.figure(figsize=(10, 6))
pollution_by_area.plot(kind='bar')
plt.title('Average Pollution Levels by
Area/City')
plt.xlabel('Area/City')
plt.ylabel('Pollution Level')
plt.show()
```

Output:

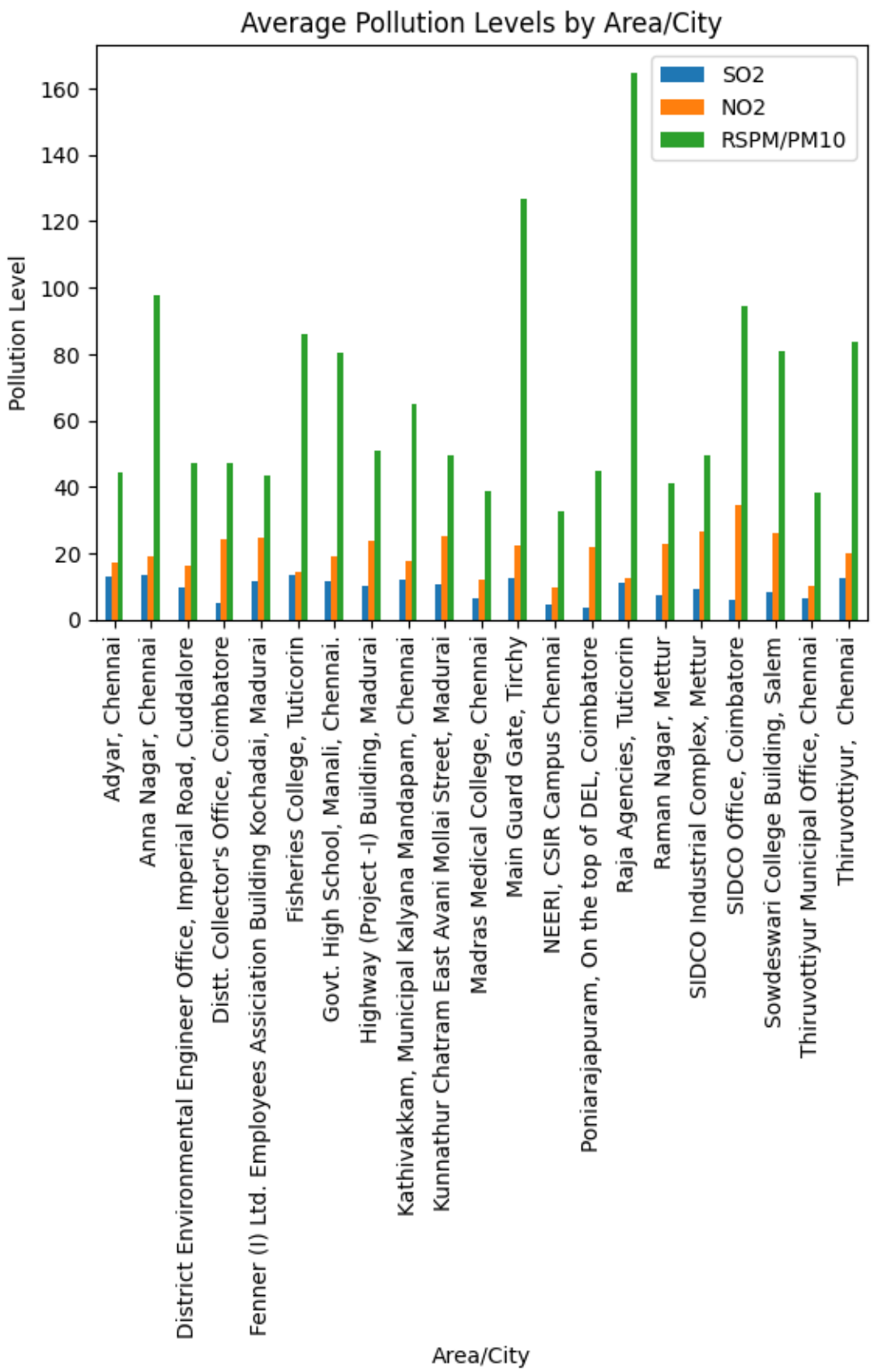


Insights:

1. Trichy has the highest RSPM Level across all region.
2. Thoothukudi has the highest SO2 level and Chennai and thoothukudi has the highest NO2 level.

Creating a bar chart to find out the highest NO2,SO2, RSPM/PM10 across Location of Monitoring Station

```
# Assuming you have data on different areas or cities, you can group the data by those areas
pollution_by_area = data.groupby('Location of Monitoring Station')['SO2', 'NO2', 'RSPM/PM10'].mean()

# Visualizing data
# Create a bar plot to show the average pollution levels in different areas
plt.figure(figsize=(10, 6))
pollution_by_area.plot(kind='bar')
plt.title('Average Pollution Levels by Area/City')
plt.xlabel('Area/City')
plt.ylabel('Pollution Level')
plt.show()
```


Average Pollution Levels by Area/City
SO2
NO2
RSPM/PM10
Pollution Level
0
20
40
60
80
100
120
140
160
Adyar, Chennai
Anna Nagar, Chennai
District Environmental Engineer Office, Imperial Road, Cuddalore
Distt. Collector's Office, Coimbatore
Fenner (I) Ltd. Employees Assiciation Building Kochadai, Madurai
Fisheries College, Tuticorin
Govt. High School, Manali, Chennai.
Highway (Project -I) Building, Madurai
Kathivakkam, Municipal Kalyana Mandapam, Chennai
Kunnathur Chatram East Avani Mollai Street, Madurai
Madras Medical College, Chennai
Main Guard Gate, Tirchy
NEERI, CSIR Campus Chennai
Poniarajapuram, On the top of DEL, Coimbatore
Raja Agencies, Tuticorin
Raman Nagar, Mettur
SIDCO Industrial Complex, Mettur
SIDCO Office, Coimbatore
Sowdeswari College Building, Salem
Thiruvottiyur Municipal Office, Chennai
Thiruvottiyur, Chennai
Area/City

Output:

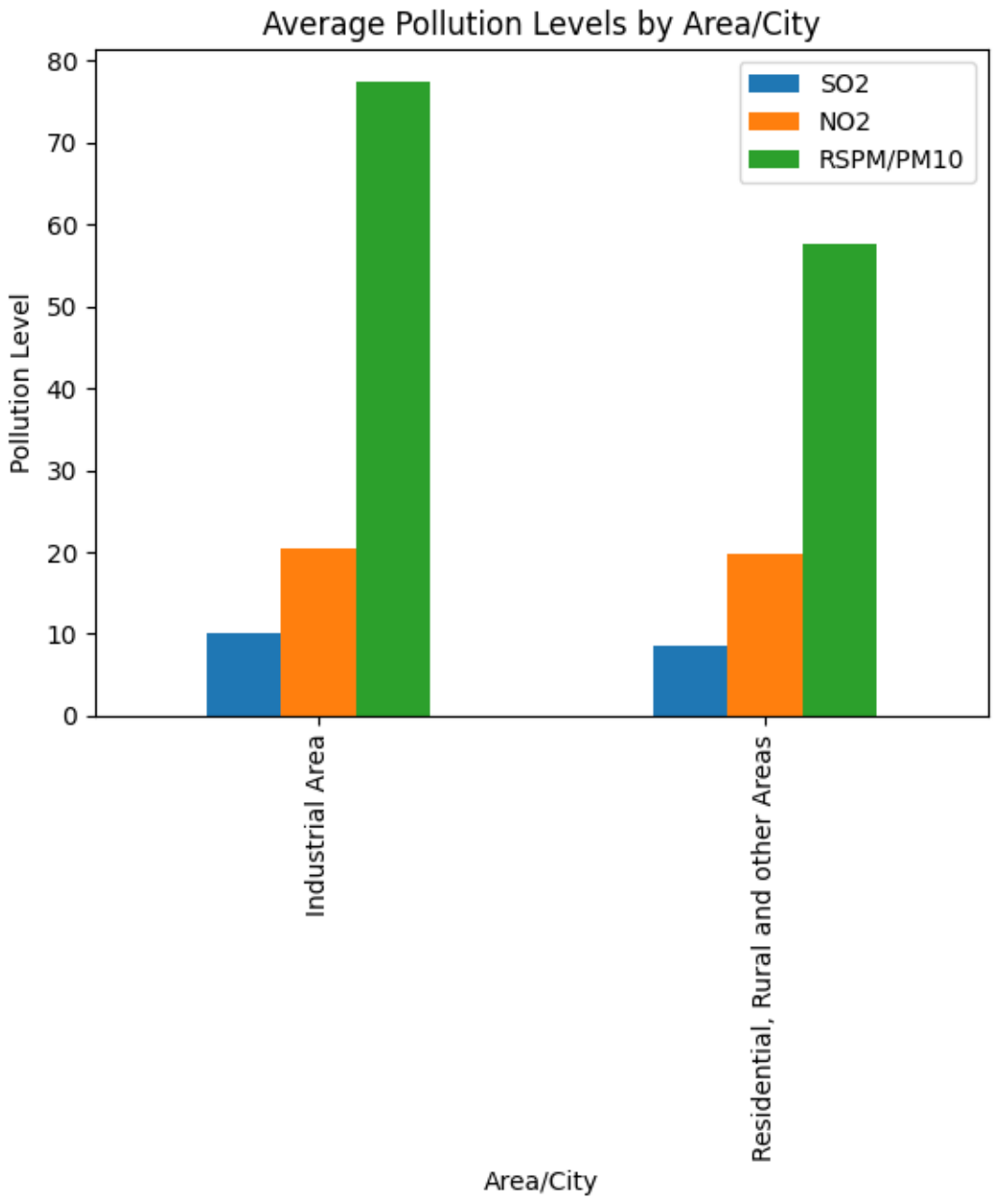
1.Raja agency of the tutcorin has the highest level of RSPM Level.

2. Adayar Agency has the highest level of So2 and No2 level.

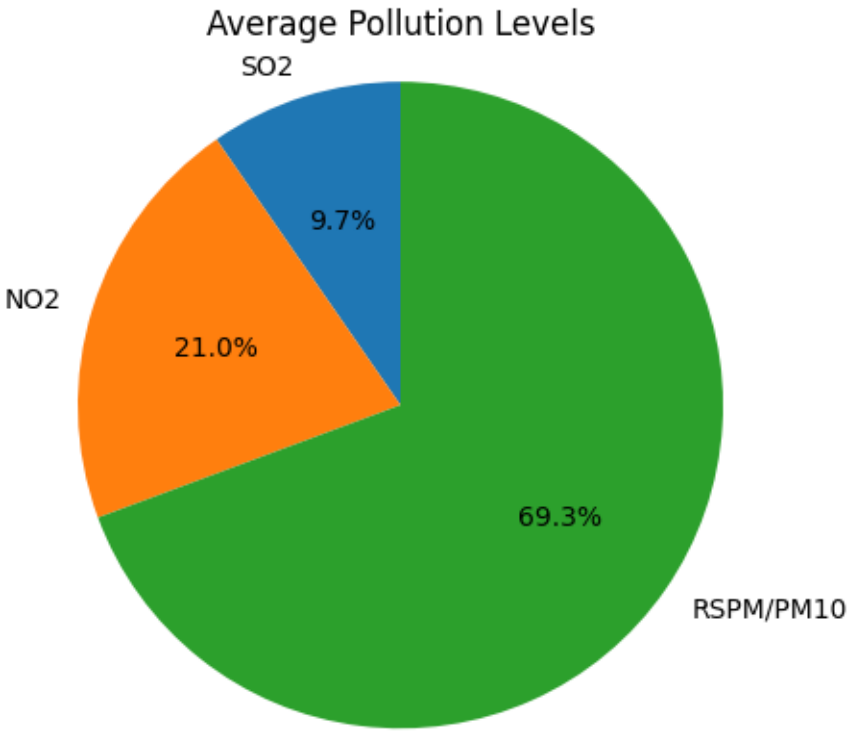
Creating a bar chart to identify the So2,No2, RSPM/PM10 across Type of Location.

```
pollution_by_area = data.groupby('Type of
Location')['SO2', 'NO2', 'RSPM/PM10'].mean()

# Visualizing data
# Create a bar plot to show the average
pollution levels in different areas
plt.figure(figsize=(10, 6))
pollution_by_area.plot(kind='bar')
plt.title('Average Pollution Levels by
Area/City')
plt.xlabel('Area/City')
plt.ylabel('Pollution Level')
plt.show()
```

Output:


Average Pollution Levels by Area/City
SO2
NO2
RSPM/PM10
Pollution Level
0
10
20
30
40
50
60
70
80
Industrial Area
Residential, Rural and other Areas
Area/City

Create a pie chart to identify the highest level of SO2,NO2,RSPM

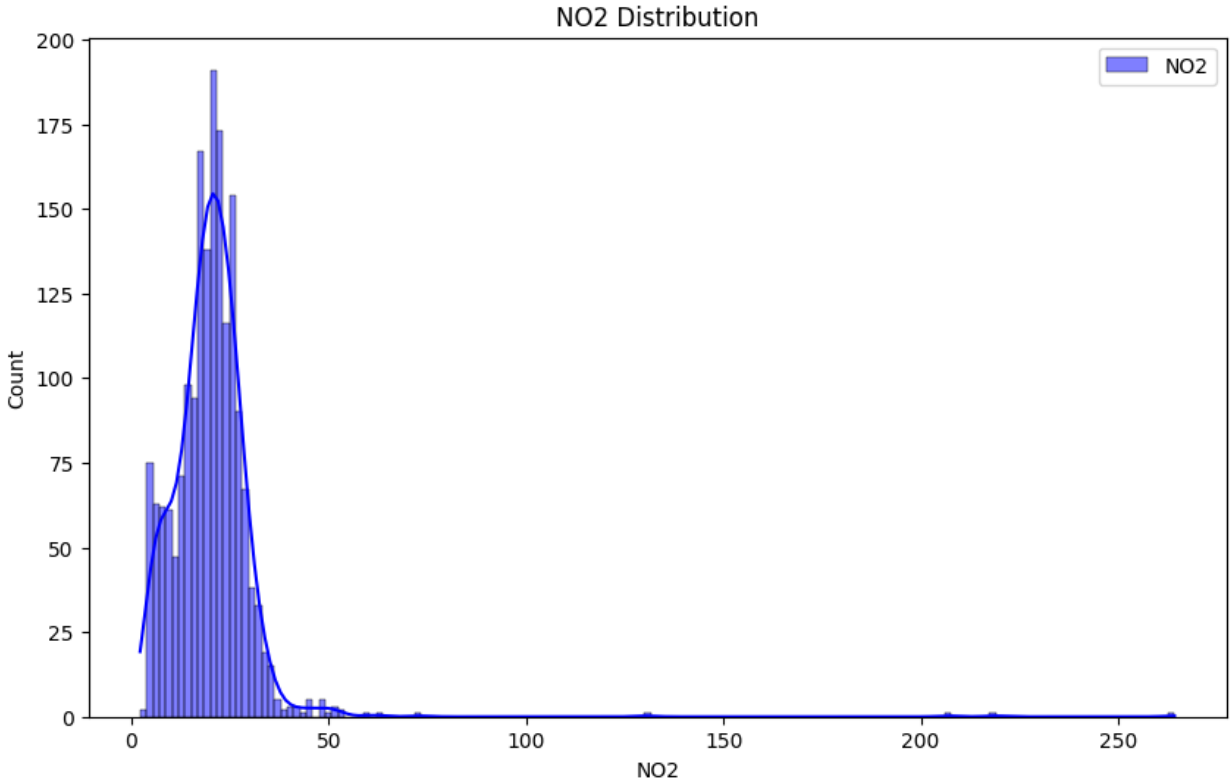


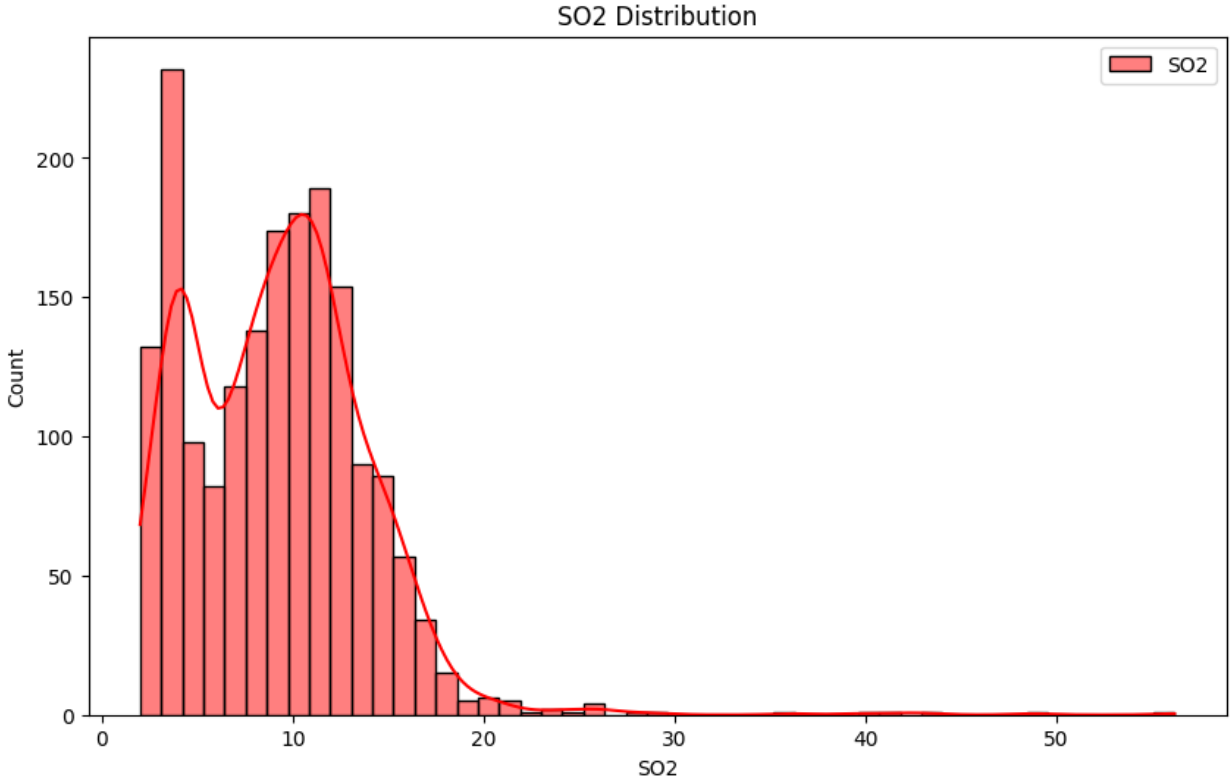
Creating a histogram for Identifying the level of NO2,SO2, RSPM/PM10.

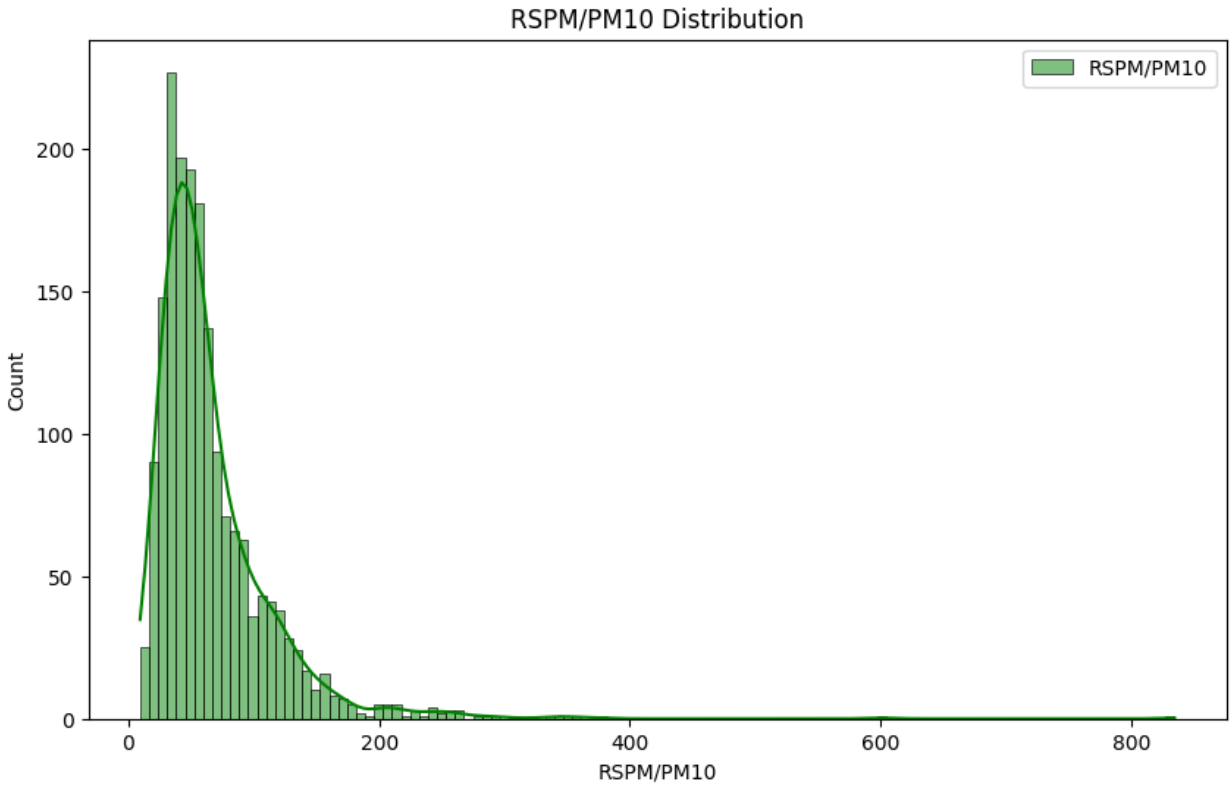
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming you have the data in a CSV file
named 'pollution_data.csv'
data = pd.read_csv('project.csv')

# Calculating average SO2, NO2, and
RSPM/PM10 levels
avg_SO2 = data['SO2'].mean()
avg_NO2 = data['NO2'].mean()
avg_RSPM_PM10 = data['RSPM/PM10'].mean()

# Printing average values
print(f"Average SO2 level: {avg_SO2}")
print(f"Average NO2 level: {avg_NO2}")
print(f"Average RSPM/PM10 level:
{avg_RSPM_PM10}")

# Identify pollution trends
# Assuming you have data on different areas
or cities, you can group the data by those
areas
```

```
# Create individual plots for each
pollutant
plt.figure(figsize=(10, 6))
sns.histplot(data['SO2'], kde=True,
color='red', label='SO2')
plt.title('SO2 Distribution')
plt.legend()
plt.show()

plt.figure(figsize=(10, 6))
sns.histplot(data['NO2'], kde=True,
color='blue', label='NO2')
plt.title('NO2 Distribution')
plt.legend()
plt.show()

plt.figure(figsize=(10, 6))
sns.histplot(data['RSPM/PM10'], kde=True,
color='green', label='RSPM/PM10')
plt.title('RSPM/PM10 Distribution')
plt.legend()
plt.show()
```

output:


NO2 Distribution
NO2
Count
NO2
0
25
50
75
100
125
150
175
200
0
50
100
150
200
250


SO2 Distribution
SO2
Count
SO2
0
50
100
150
200
0
10
20
30
40
50

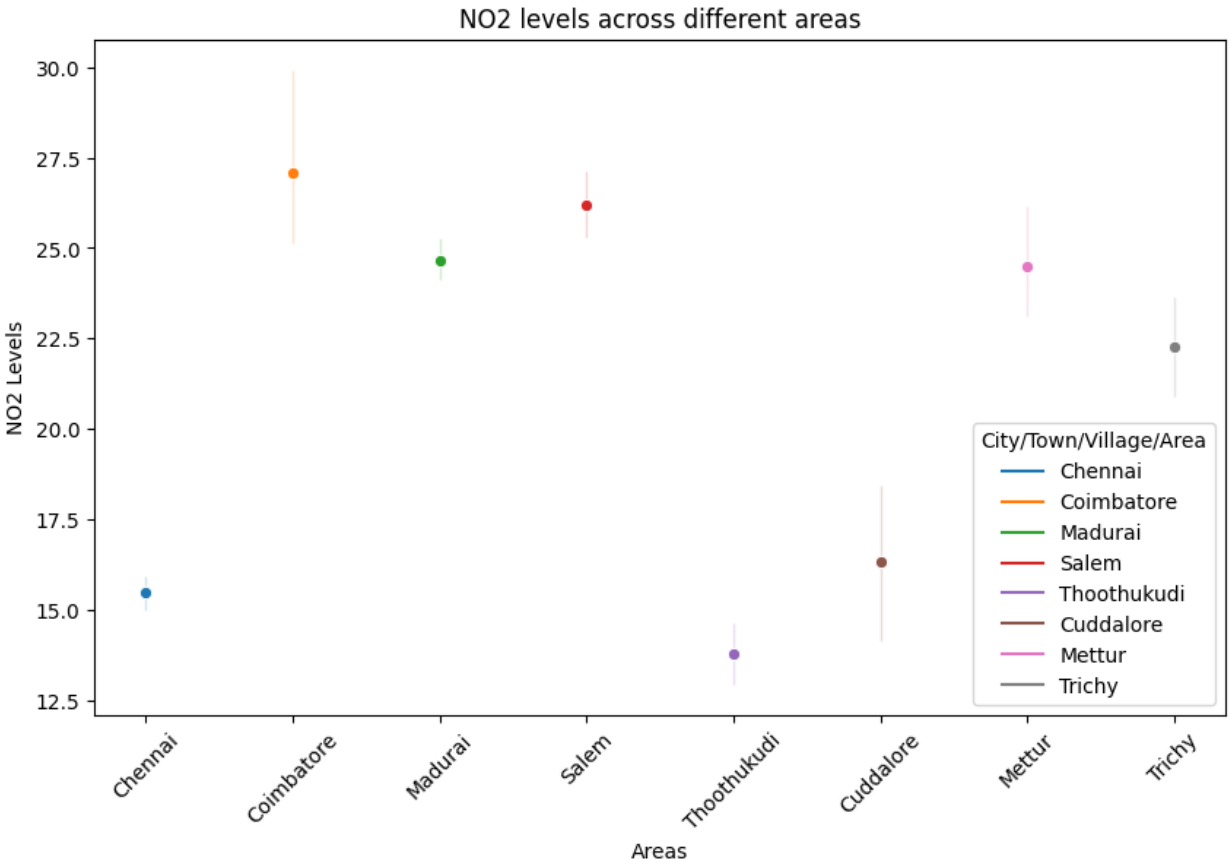


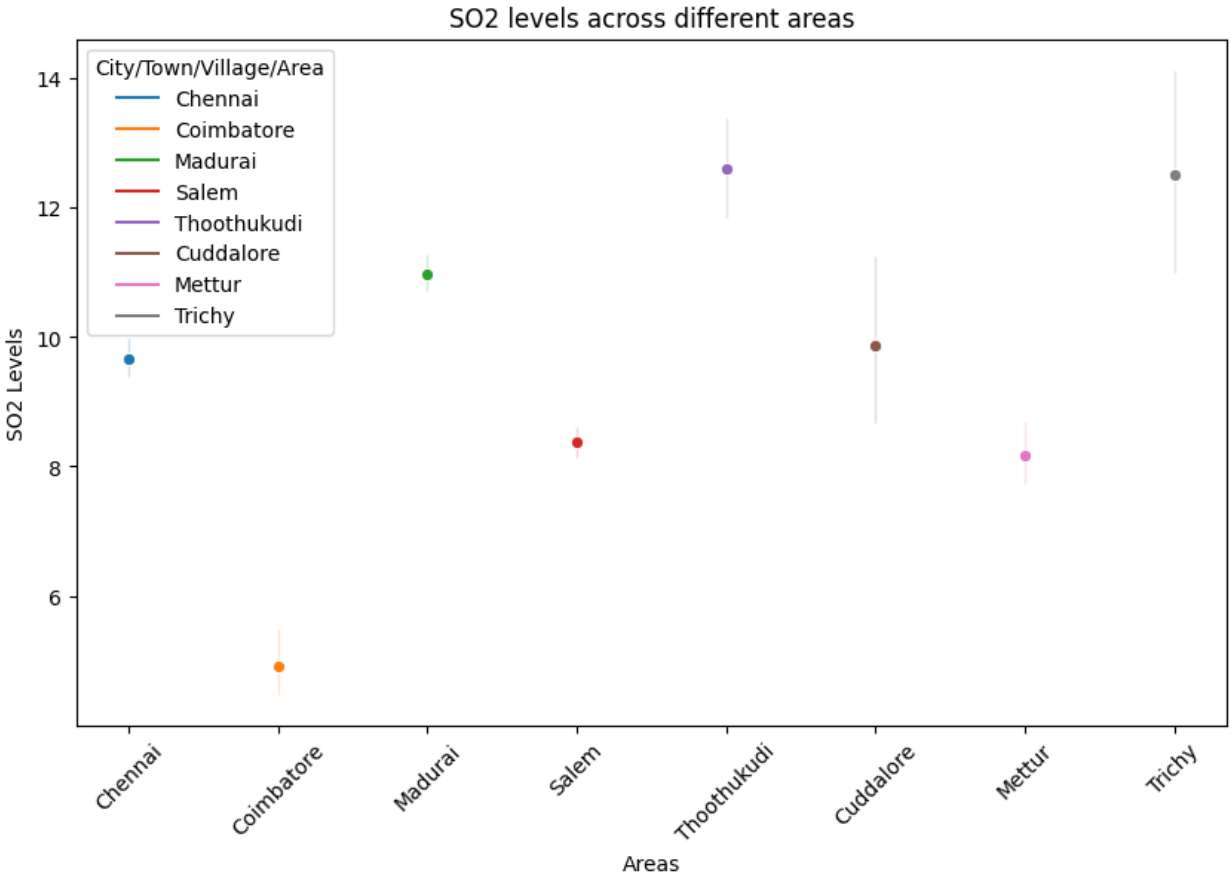
Creating a line plot to identify the highest NO2, SO2,RSPM/10 highest range of City/Town/Village/Area

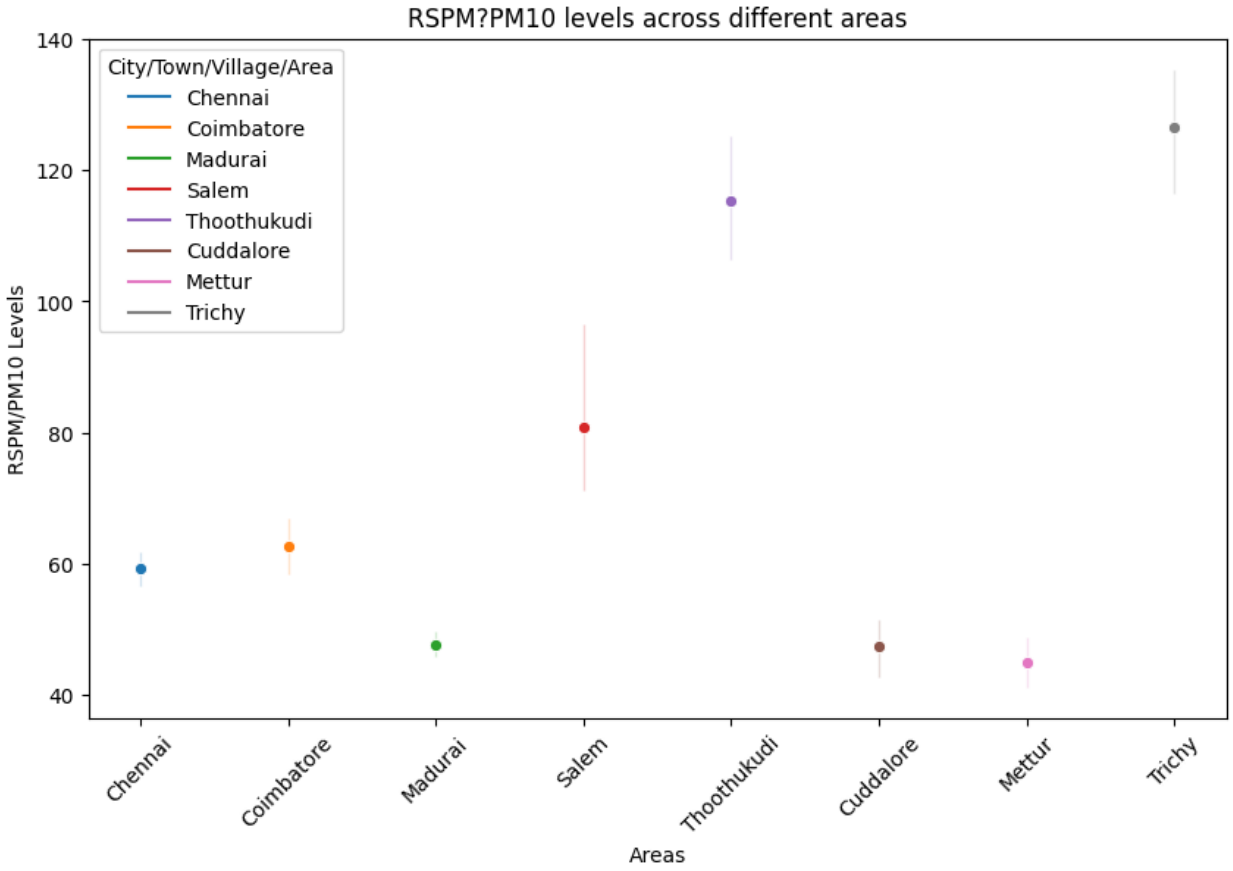
```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
df=pd.read_csv("project.csv")
plt.figure(figsize=(10, 6))
sns.lineplot(data=df,
x='City/Town/Village/Area', y='NO2',
marker='o', hue='City/Town/Village/Area')

plt.title('NO2 levels across different
areas')
plt.xlabel('Areas')
```

```
plt.ylabel('NO2 Levels')
plt.xticks(rotation=45)
plt.show()
sns.lineplot(data=df,
x='City/Town/Village/Area', y='SO2',
marker='o', hue='City/Town/Village/Area')
plt.title('SO2 levels across different
areas')
plt.xlabel('Areas')
plt.ylabel('SO2 Levels')
plt.xticks(rotation=45)
plt.show()
sns.lineplot(data=df,
x='City/Town/Village/Area', y='RSPM/PM10',
marker='o', hue='City/Town/Village/Area')
plt.title('RSPM?PM10 levels across
different areas')
plt.xlabel('Areas')
plt.ylabel('RSPM?PM10 Levels')
plt.xticks(rotation=45)
```

OUTPUT:


NO2 levels across different areas
NO2 Levels
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy
Areas
City/Town/Village/Area
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy


SO2 levels across different areas
City/Town/Village/Area
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy
SO2 Levels
14
12
10
8
6
Areas
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy

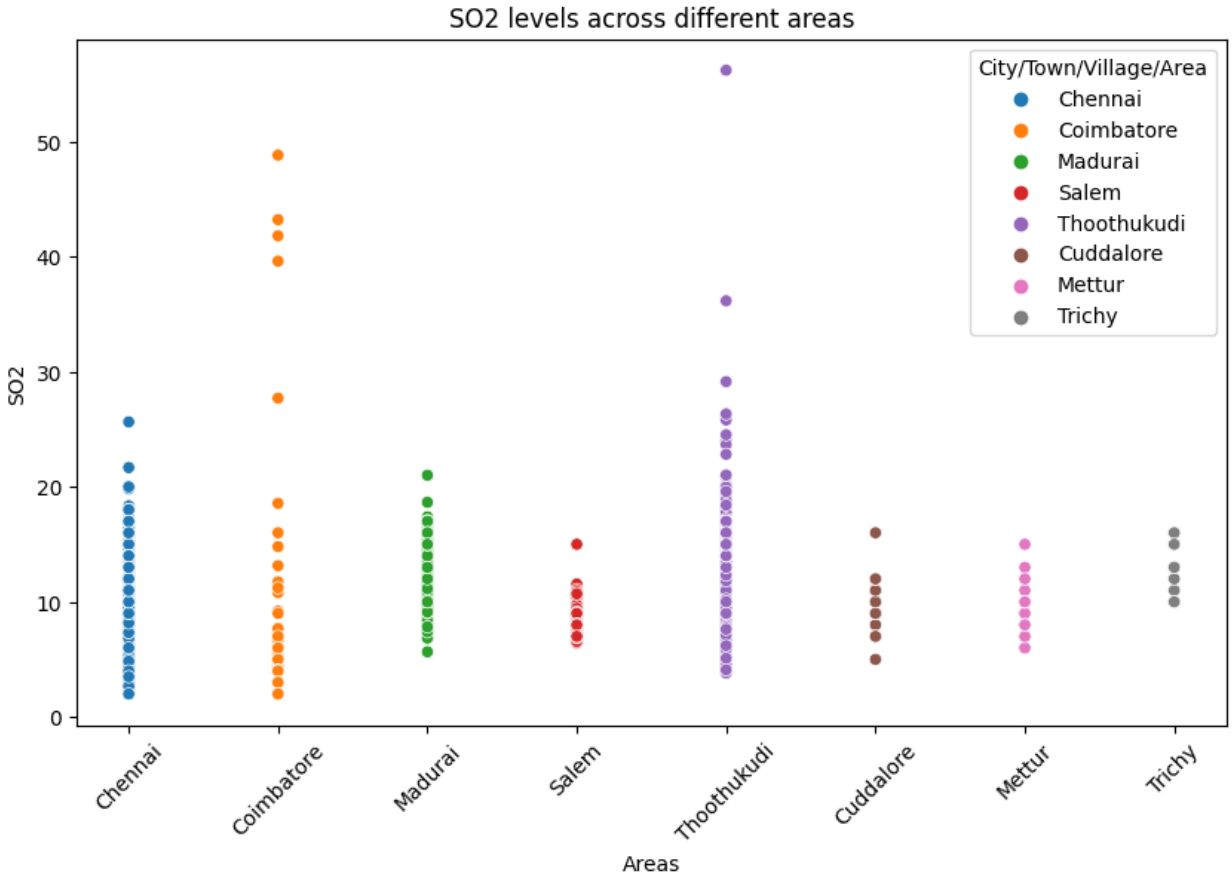


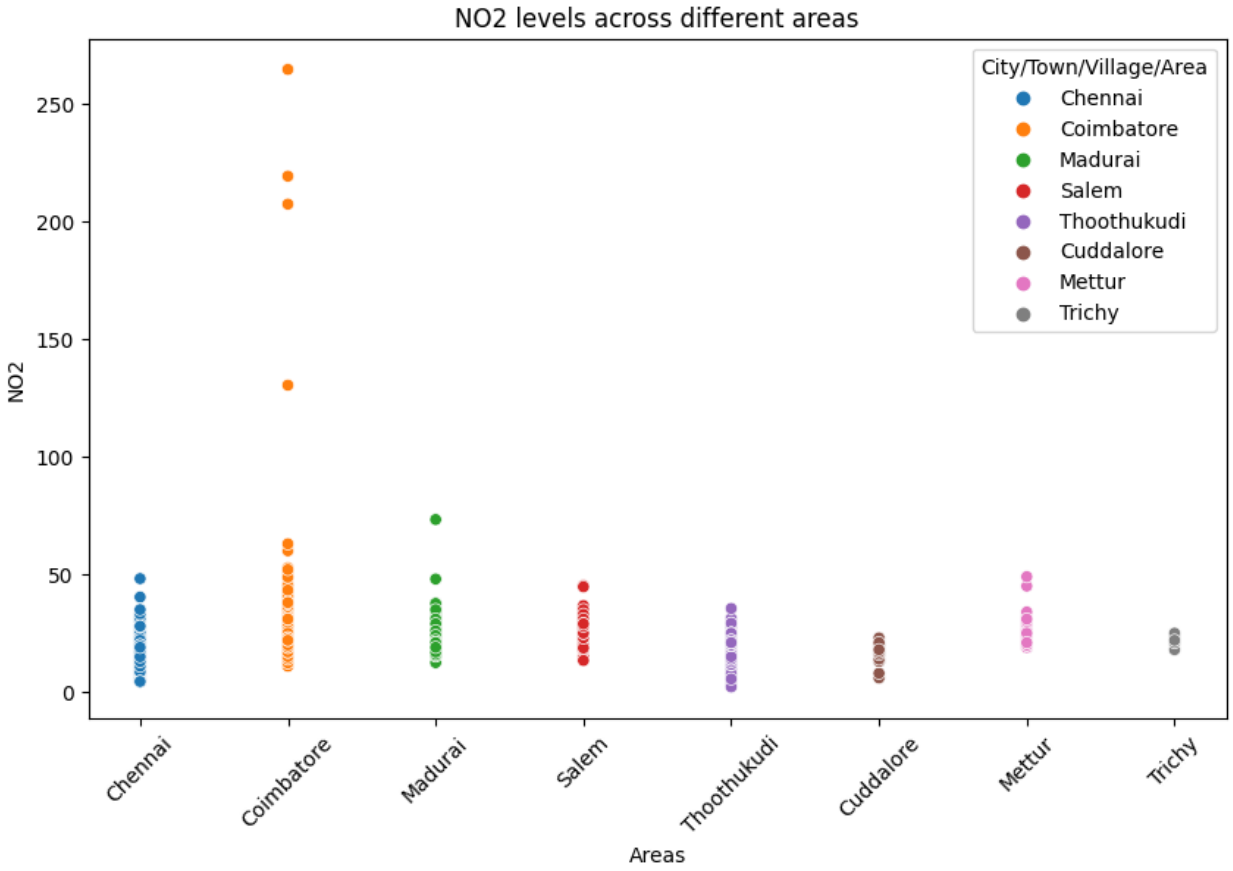
Creating a Scatter plot for Identifying the level of NO2,SO2, RSPM/PM10

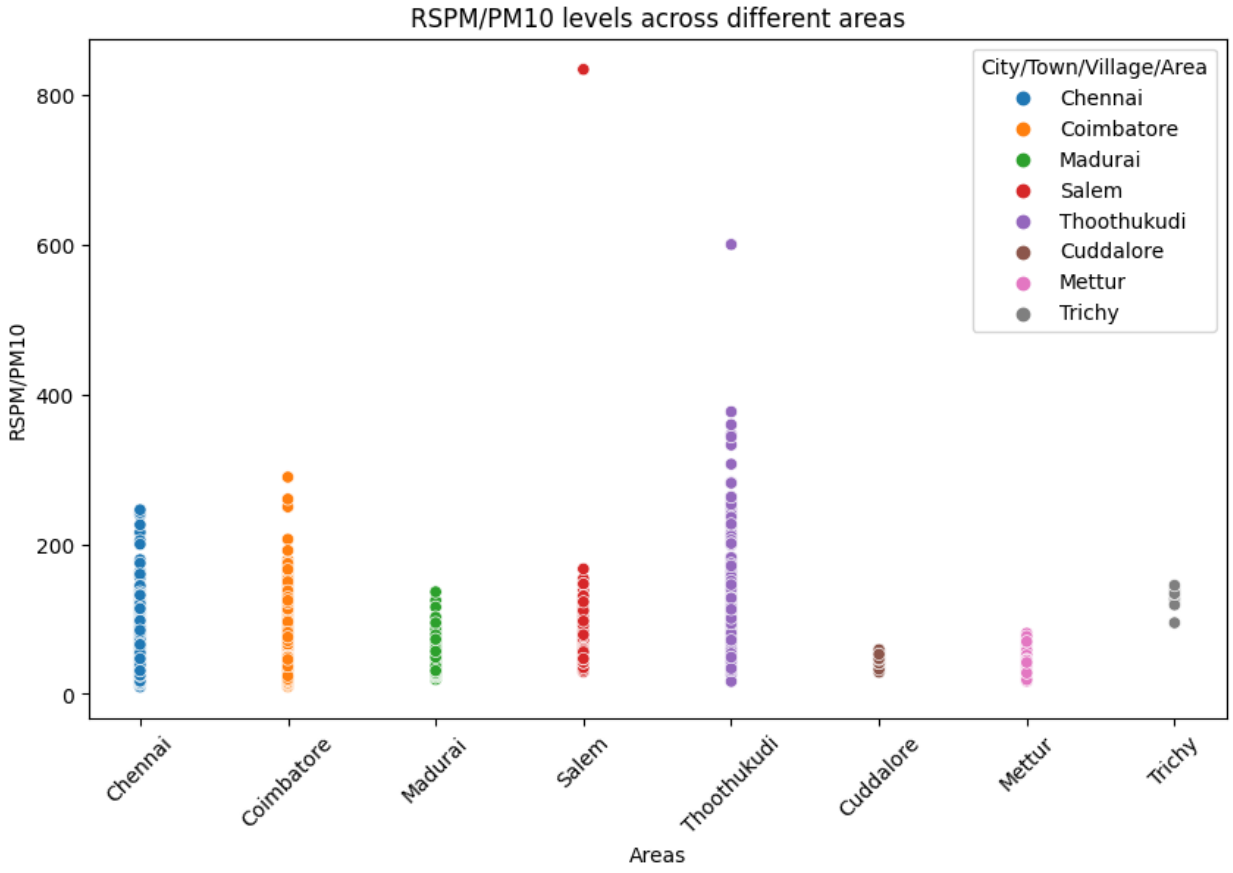
```
 import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
df=pd.read_csv("project.csv")
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df,
x='City/Town/Village/Area', y='SO2', marker='o',
hue='City/Town/Village/Area')

plt.title('SO2 levels across different areas')
plt.xlabel('Areas')
plt.ylabel('SO2')
```

```
plt.xticks(rotation=45)
plt.show()
sns.scatterplot(data=df,
x='City/Town/Village/Area', y='NO2', marker='o',
hue='City/Town/Village/Area')

plt.title('NO2 levels across different areas')
plt.xlabel('Areas')
plt.ylabel('NO2')
plt.xticks(rotation=45)
plt.show()
sns.scatterplot(data=df,
x='City/Town/Village/Area', y='SO2', marker='o',
hue='City/Town/Village/Area')

plt.title('RSPM/PM10 levels across different
areas')
plt.xlabel('Areas')
plt.ylabel('RSPM/PM10')
plt.xticks(rotation=45)
plt.show()
```

output:

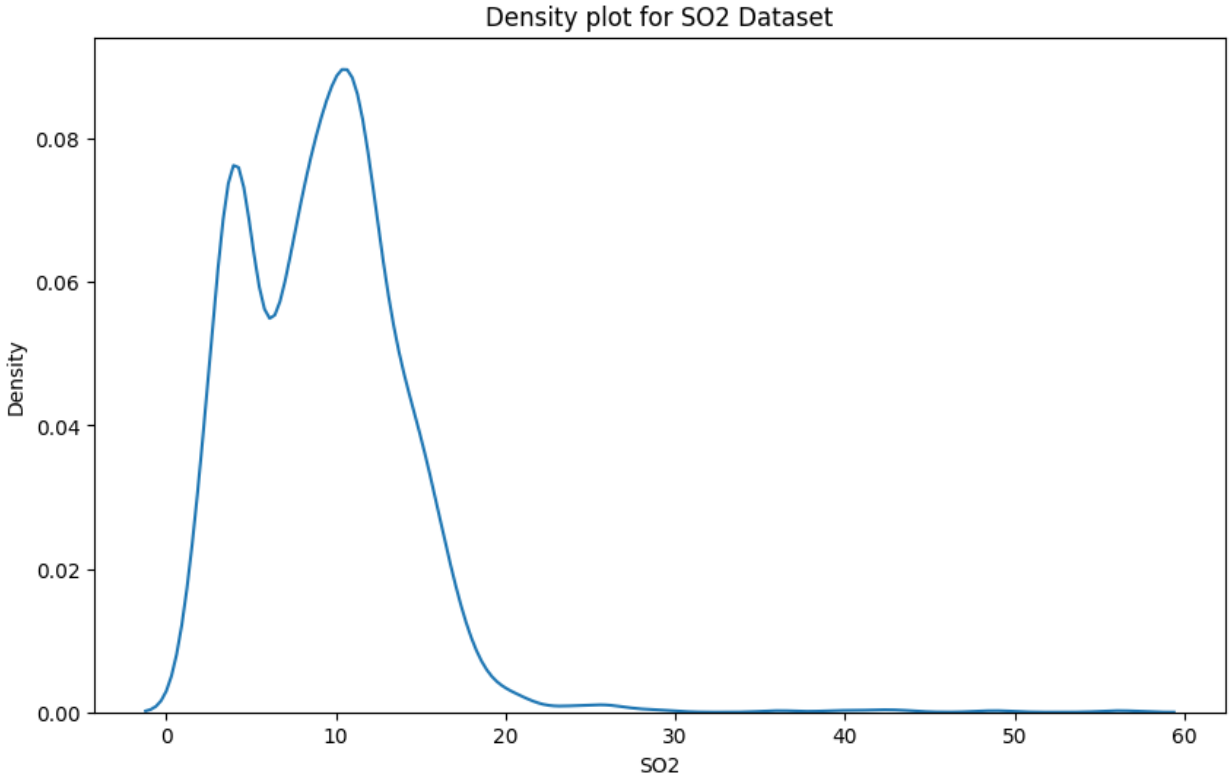

SO2 levels across different areas
City/Town/Village/Area
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy
SO2
0
10
20
30
40
50
Areas
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy

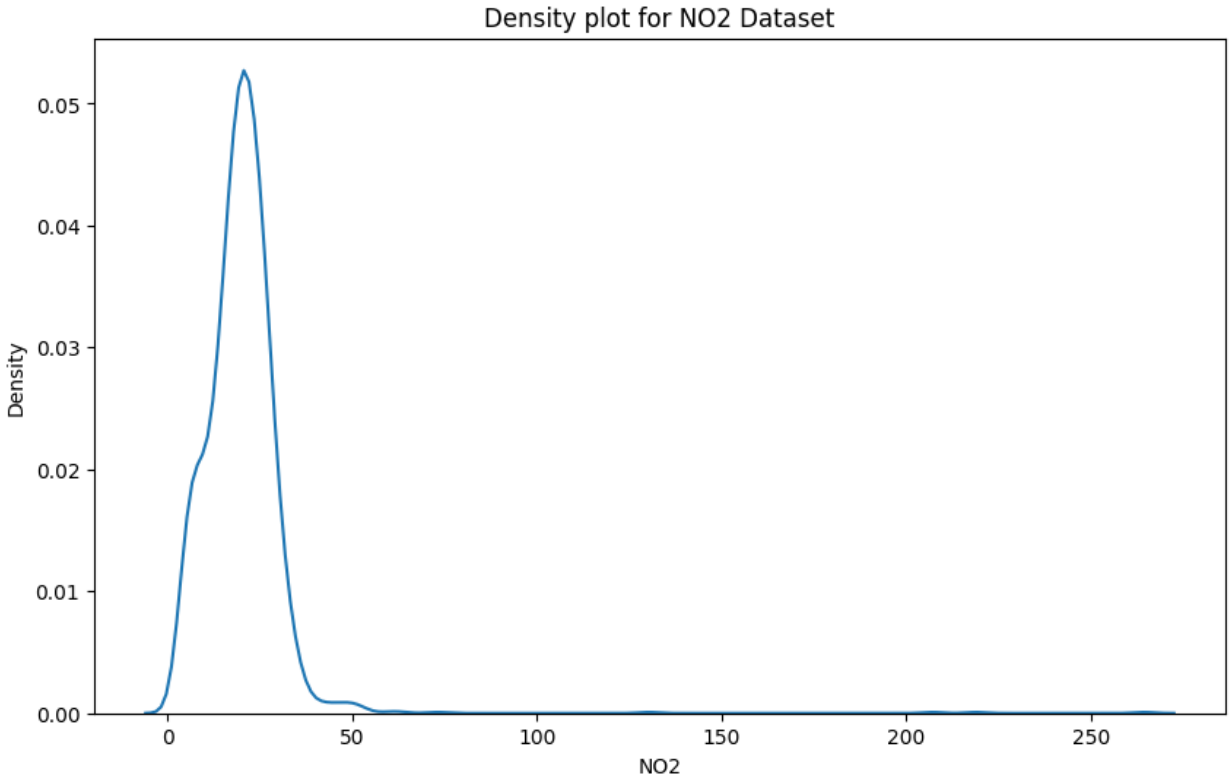

NO2 levels across different areas
City/Town/Village/Area
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy
NO2
0
50
100
150
200
250
Areas
Chennai
Coimbatore
Madurai
Salem
Thoothukudi
Cuddalore
Mettur
Trichy

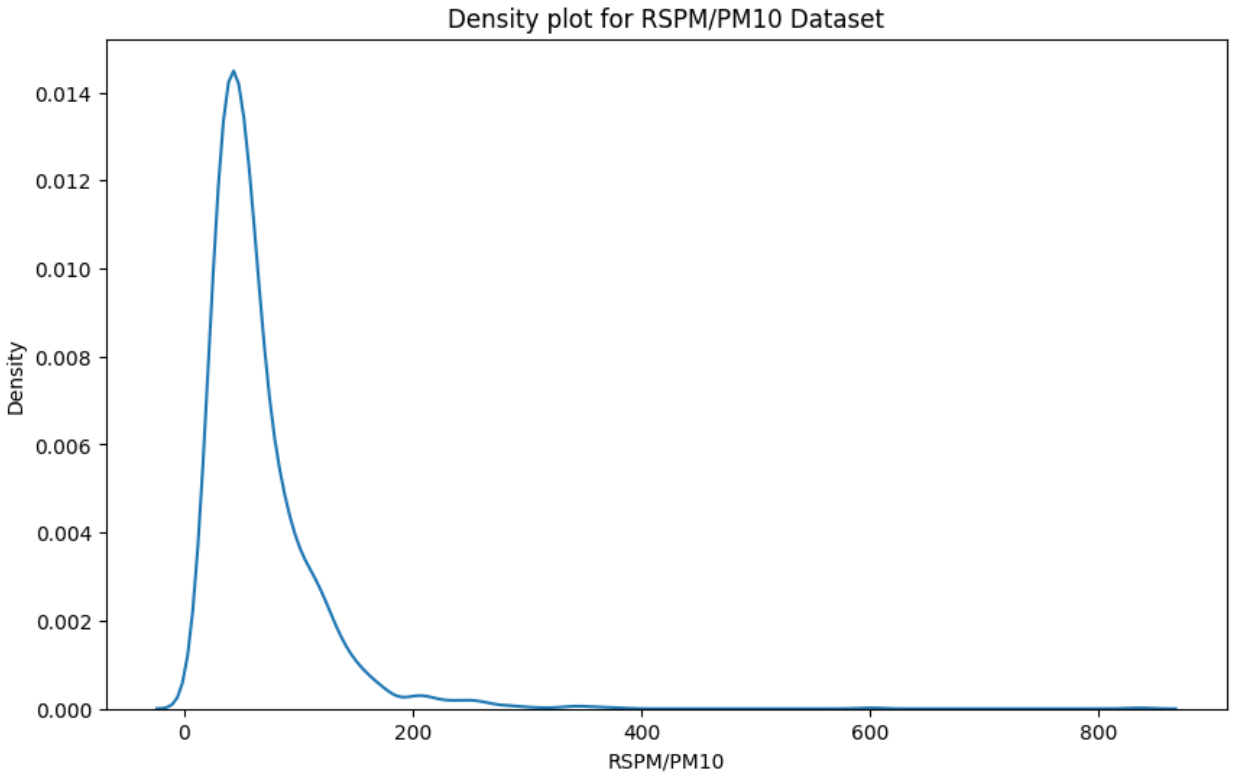


Creating a Distance plot to identify the highest NO2, SO2,RSPM/10 highest range of City/Town/Village/Area

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
df=pd.read_csv("project.csv")
plt.figure(figsize=(10, 6))
x=df["SO2"]
sns.distplot(x,hist=False)
plt.title("Density plot for SO2 Dataset")
plt.show()
```


Density plot for SO2 Dataset
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
50
60
SO2

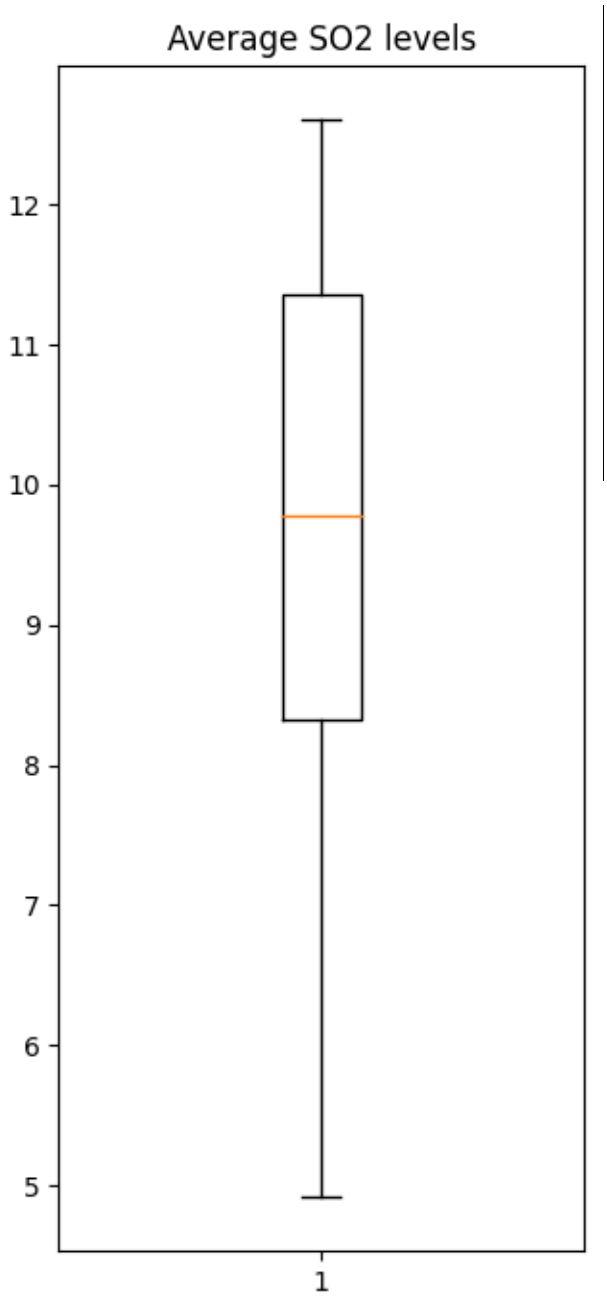

Density plot for NO2 Dataset
Density
0.00
0.01
0.02
0.03
0.04
0.05
0
50
100
150
200
250
NO2

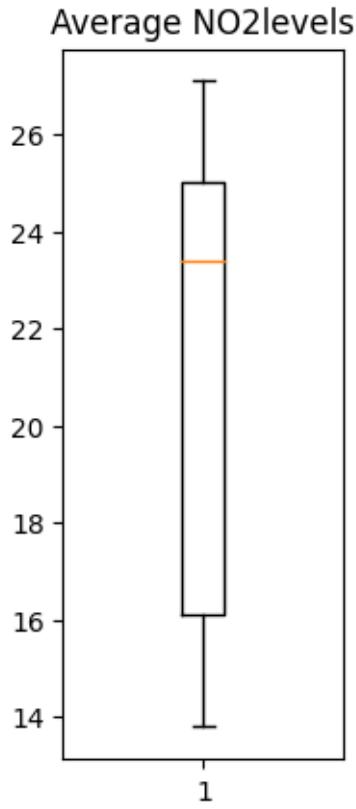

Density plot for RSPM/PM10 Dataset
Density
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0
200
400
600
800
RSPM/PM10

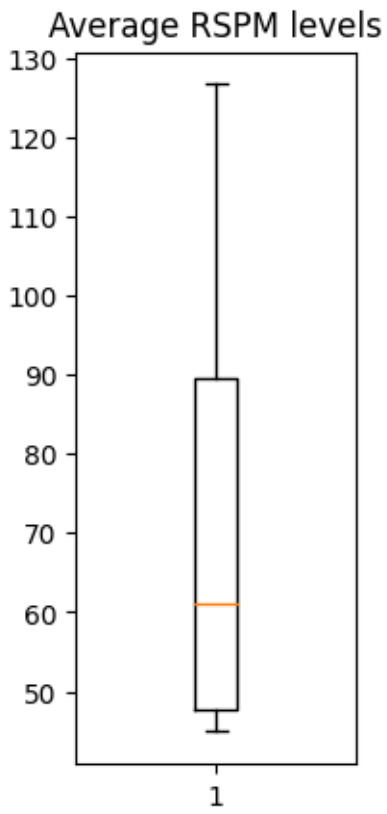
Creating a box plot

```
 import pandas as pd
import matplotlib.pyplot as plt

# Load the data from the CSV file
data = pd.read_csv('project.csv')  # Update 'project.csv' with the path to your CSV file

# Assuming the data has columns for 'City/Town/Village/Area', 'SO2', 'NO2', 'RSPM/PM10', and the necessary data for the box plot

# Calculate the average values of SO2, NO2, and RSPM/PM10 across different areas
avg_data = data.groupby('City/Town/Village/Area')[['SO2', 'NO2', 'RSPM/PM10']].mean()

# Create box plots for the three pollutants
plt.figure(figsize=(12, 8))
plt.subplot(1, 3, 1)
plt.boxplot(avg_data['SO2'])
plt.title('Average SO2 levels')

plt.subplot(1, 3, 2)
plt.boxplot(avg_data['NO2'])
plt.title('Average NO2 levels')

plt.subplot(1, 3, 3)
plt.boxplot(avg_data['RSPM/PM10'])
plt.title('Average RSPM/PM10 levels')

plt.show()
```

output:


Average SO2 levels
12
11
10
9
8
7
6
5
1


Average NO2levels
26
24
22
20
18
16
14
1

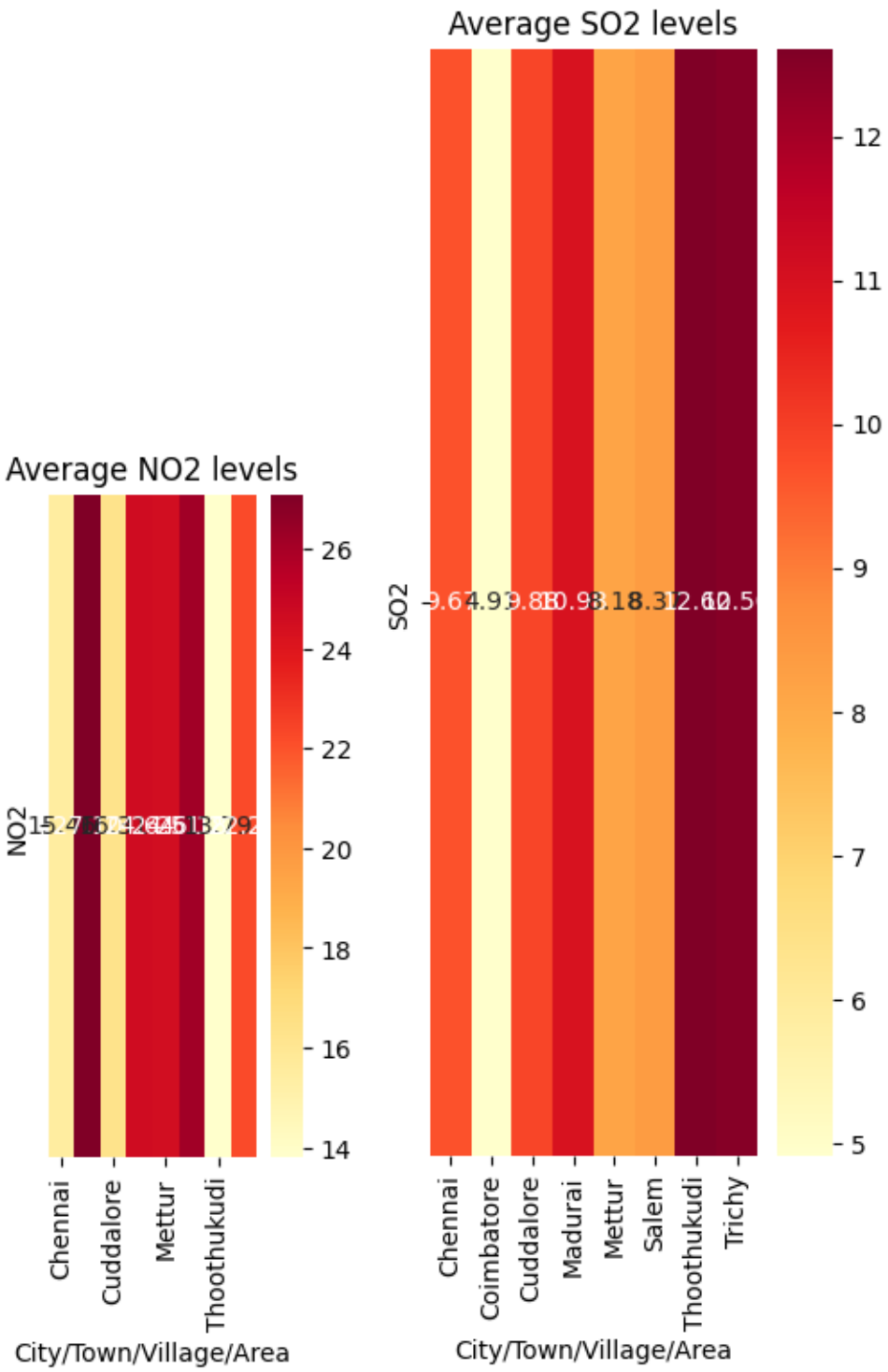


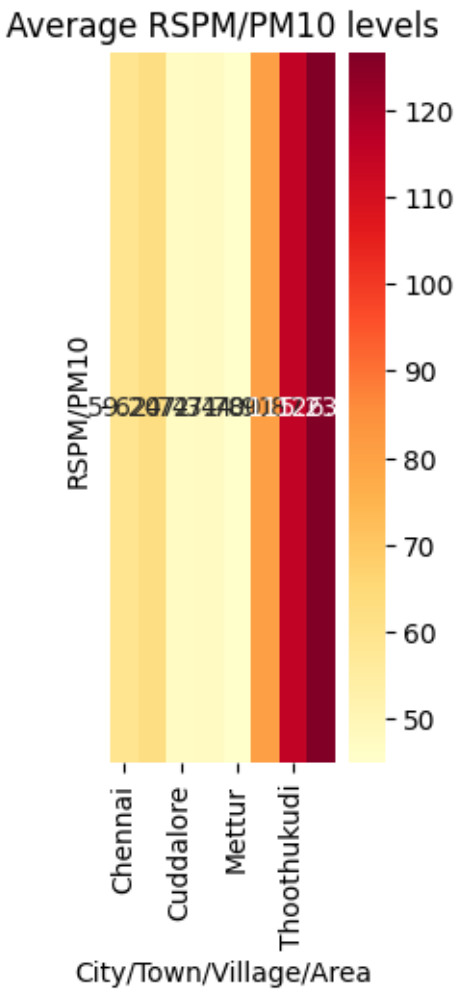
Create a heat map to find out highest range of No2 ,SO2,RSPM/PM10

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the data from the CSV file
data = pd.read_csv('project.csv')  # Update 'project.csv' with the path to your CSV file

# Assuming the data has columns for 'City/Town/Village/Area', 'SO2', 'NO2',
```

```
'RSPM/PM10', and the necessary data for the
heatmap

# Calculate the average values of SO2, NO2,
and RSPM/PM10 across different areas
avg_data =
data.groupby('City/Town/Village/Area')[['SO
2', 'NO2', 'RSPM/PM10']].mean()

# Create a heatmap for the three pollutants
plt.figure(figsize=(10, 8))
plt.subplot(1, 3, 1)
sns.heatmap(avg_data[['SO2']].transpose(),
cmap='YlOrRd', annot=True, fmt=".2f")
plt.title('Average SO2 levels')

plt.subplot(1, 3, 2)
sns.heatmap(avg_data[['NO2']].transpose(),
cmap='YlOrRd', annot=True, fmt=".2f")
plt.title('Average NO2 levels')

plt.subplot(1, 3, 3)
sns.heatmap(avg_data[['RSPM/PM10']].transpo
se(), cmap='YlOrRd', annot=True, fmt=".2f")
plt.title('Average RSPM/PM10 levels')

plt.show()
```

Output:


Average NO2 levels
NO2
26
24
22
20
18
16
14
Chennai
Cuddalore
Mettur
Thoothukudi
City/Town/Village/Area
Average SO2 levels
SO2
12
11
10
9
8
7
6
5
Chennai
Coimbatore
Cuddalore
Madurai
Mettur
Salem
Thoothukudi
Trichy
City/Town/Village/Area

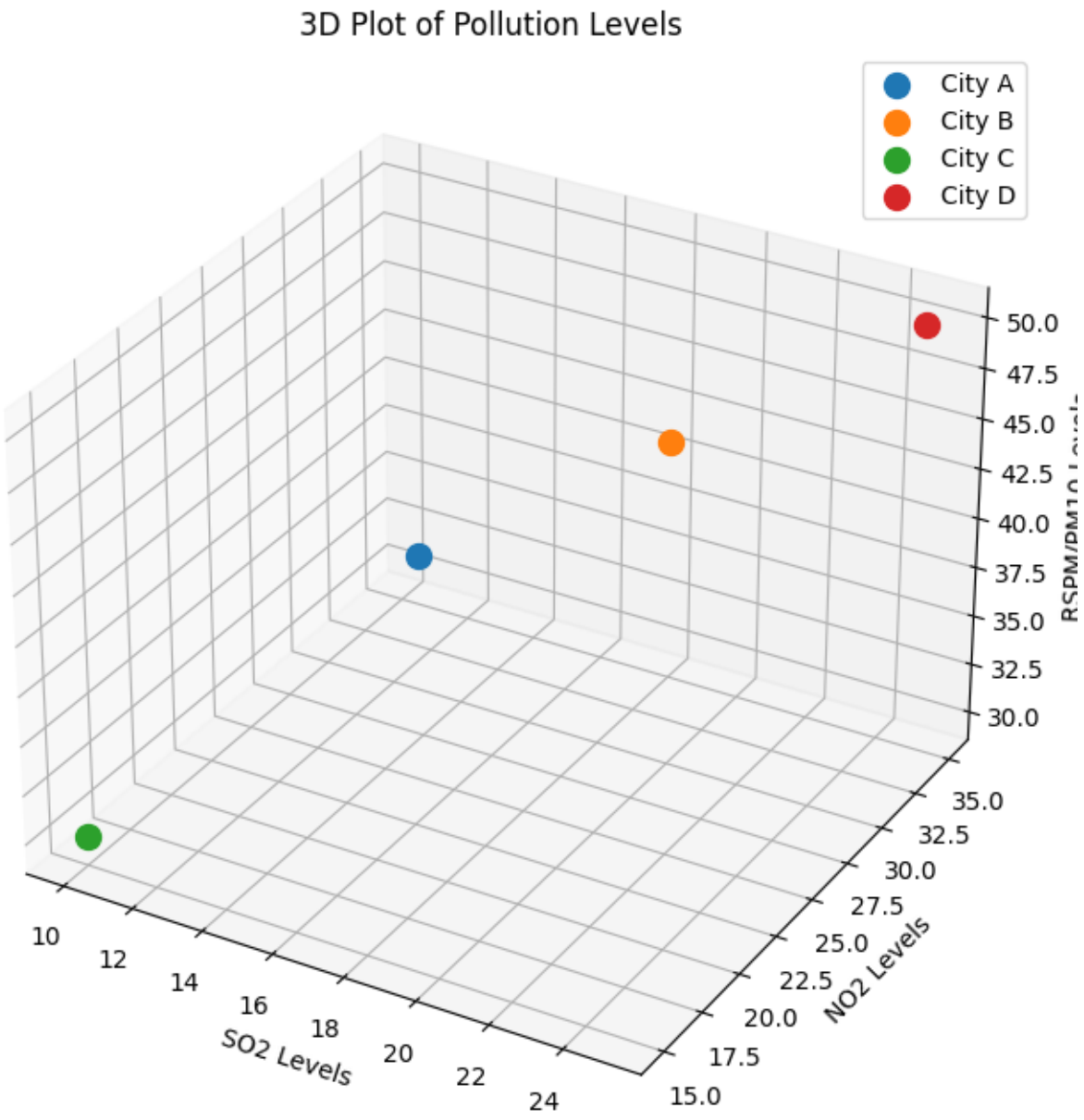


Create a 3D plot identify highest level of NO2 and So2 plotted against the city/Town/village /area:

```
mport numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Sample data for demonstration

# Extracting the data
cities = data['City/Town/Village/Area']
so2_levels = data['SO2']
```

```
no2_levels = data['NO2']
rspm_levels = data['RSPM/PM10']

# Creating a 3D plot
fig = plt.figure(figsize=(10, 8))
ax = fig.add_subplot(111, projection='3d')

# Plotting the data
for i in range(len(cities)):
    ax.scatter(so2_levels[i],
no2_levels[i], rspm_levels[i], marker='o',
s=100, label=cities[i])

# Adding labels and title
ax.set_xlabel('SO2 Levels')
ax.set_ylabel('NO2 Levels')
ax.set_zlabel('RSPM/PM10 Levels')
ax.set_title('3D Plot of Pollution Levels')

# Adding a legend
ax.legend()

# Show the plot
```

Output:


3D Plot of Pollution Levels
City A
City B
City C
City D
50.0
47.5
45.0
42.5
40.0
37.5
35.0
32.5
30.0
RSPM/PM10 Levels
35.0
32.5
30.0
27.5
25.0
22.5
20.0
17.5
15.0
NO2 Levels
10
12
14
16
18
20
22
24
SO2 Levels