

## Phase 4 : Development Part 4

Aim:

To Calculate the average SO<sub>2</sub>,NO<sub>2</sub>,RSPM/ PM<sub>10</sub> Levels across different monitoring stations ,cities, States.To identify the pollution trends and areas With high pollution levels.To create visualizatio Using Data Visualization libraries like Seaborn , Matplotlib,etc.

1.Python code:

```
import pandas as pd
df=pd.read_csv("C:\User\admin\Desktop\ibm\air.csv")
df['SO2']=df['SO2'].interpolate()
df['NO2']=df['NO2'].interpolate()
df=df.drop(columns=df.columns[-1], axis=1,inplace=False)
print(df)
df.describe()
```

Output:

```
Untitled1.ipynb x Untitled2.ipynb x +
+ ✂ 📄 ▶ ■ ↺ ⏩ Code ⌵ ☰

2874 Central Bus Stand, Trichy
2875 Central Bus Stand, Trichy
2876 Central Bus Stand, Trichy
2877 Central Bus Stand, Trichy
2878 Central Bus Stand, Trichy

Agency \
0 Tamilnadu State Pollution Control Board
1 Tamilnadu State Pollution Control Board
2 Tamilnadu State Pollution Control Board
3 Tamilnadu State Pollution Control Board
4 Tamilnadu State Pollution Control Board
...
2874 Tamilnadu State Pollution Control Board
2875 Tamilnadu State Pollution Control Board
2876 Tamilnadu State Pollution Control Board
2877 Tamilnadu State Pollution Control Board
2878 Tamilnadu State Pollution Control Board

Type of Location SO2 NO2 RSPM/PM10
0 Industrial Area 11.0 17.0 55.0
1 Industrial Area 13.0 17.0 45.0
2 Industrial Area 12.0 18.0 50.0
3 Industrial Area 15.0 16.0 46.0
4 Industrial Area 13.0 14.0 42.0
...
2874 Residential, Rural and other Areas 15.0 18.0 102.0
2875 Residential, Rural and other Areas 12.0 14.0 91.0
2876 Residential, Rural and other Areas 19.0 22.0 100.0
2877 Residential, Rural and other Areas 15.0 17.0 95.0
2878 Residential, Rural and other Areas 14.0 16.0 94.0

[2879 rows x 10 columns]

[77]:
```

	Stn Code	SO2	NO2	RSPM/PM10
count	2879.000000	2879.000000	2879.000000	2875.000000
mean	475.750261	11.510247	22.136506	62.494261
std	277.675577	5.051316	7.126277	31.368745
min	38.000000	2.000000	5.000000	12.000000
25%	238.000000	8.000000	17.000000	41.000000
50%	366.000000	12.000000	22.000000	55.000000
75%	764.000000	15.000000	25.000000	78.000000

## 2. Finding average for SO<sub>2</sub>, NO<sub>2</sub>, RSPM/PM<sub>10</sub> on different Locations.

Code:

```
#Average levels in chennai
chennai = df.loc[df['City/Town/Village/Area'] == 'Chennai']
avg_so2_chenn=chennai['SO2'].mean()
avg_no2_chenn=chennai['NO2'].mean()
avg_rspm_chenn=chennai['RSPM/PM10'].mean()
print(avg_so2_chenn)
print(avg_no2_chenn)
print(avg_rspm_chenn)
```

Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
chennai = df.loc[df['City/Town/Village/Area'] == 'Chennai']
avg_so2_chenn=chennai['SO2'].mean()
avg_no2_chenn=chennai['NO2'].mean()
avg_rspm_chenn=chennai['RSPM/PM10'].mean()
print("Average SO2 in chennai",avg_so2_chenn)
print("Average NO2 in chennai",avg_no2_chenn)
print("Average RSPM in chennai",avg_rspm_chenn)
```

```
Average SO2 in chennai 13.025
Average NO2 in chennai 22.1035
Average RSPM in chennai 58.998
```

#Average levels in Trichy:

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
Trichy = df.loc[df['City/Town/Village/Area'] == 'Trichy']
```

```
avg_so2_tri=Trichy['SO2'].mean()
```

```
avg_no2_tri=Trichy['NO2'].mean()
```

```
avg_rspm_tri=Trichy['RSPM/PM10'].mean()
```

```
print("Average SO2 in Trichy",avg_so2_tri)
```

```
print("Average NO2 in Trichy",avg_no2_tri)
```

```
print("Average RSPM in Trichy",avg_rspm_tri)
```

Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
Trichy = df.loc[df['City/Town/Village/Area'] == 'Trichy']
avg_so2_tri=Trichy['SO2'].mean()
avg_no2_tri=Trichy['NO2'].mean()
avg_rspm_tri=Trichy['RSPM/PM10'].mean()
print("Average SO2 in Trichy",avg_so2_tri)
print("Average NO2 in Trichy",avg_no2_tri)
print("Average RSPM in Trichy",avg_rspm_tri)
```

```
Average SO2 in Trichy 15.279291553133515
```

```
Average NO2 in Trichy 18.682561307901906
```

```
Average RSPM in Trichy 85.05449591280654
```

```
#Average levels in Coimbatore
import matplotlib.pyplot as plt
import seaborn as sns
coim = df.loc[df['City/Town/Village/Area'] == 'Coimbatore']
avg_so2_coi=coim['SO2'].mean()
avg_no2_coi=coim['NO2'].mean()
avg_rspm_coi=coim['RSPM/PM10'].mean()
print("Average SO2 in Coimbatore ",avg_so2_coi)
print("Average NO2 in Coimbatore",avg_no2_coi)
print("Average RSPM in Coimbatore",avg_rspm_coi)
```

Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
coim = df.loc[df['City/Town/Village/Area'] == 'Coimbatore']
avg_so2_coi=coim['SO2'].mean()
avg_no2_coi=coim['NO2'].mean()
avg_rspm_coi=coim['RSPM/PM10'].mean()
print("Average SO2 in Coimbatore ",avg_so2_coi)
print("Average NO2 in Coimbatore",avg_no2_coi)
print("Average RSPM in Coimbatore",avg_rspm_coi)
```

```
Average SO2 in Coimbatore 4.546075085324232
Average NO2 in Coimbatore 25.339590443686006
Average RSPM in Coimbatore 49.217241379310344
```

```
#Average levels in Mettur
import matplotlib.pyplot as plt
import seaborn as sns
mettur= df.loc[df['City/Town/Village/Area'] ==
'Mettur']
avg_so2_mett=mettur['SO2'].mean()
avg_no2_mett=mettur['NO2'].mean()
avg_rspm_mett=mettur['RSPM/PM10'].mean()
print("Average SO2 in Mettur ",avg_so2_mett)
print("Average NO2 in Mettur",avg_no2_mett)
print("Average RSPM in Mettur",avg_rspm_mett)
```

Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
mettur= df.loc[df['City/Town/Village/Area'] == 'Mettur']
avg_so2_mett=mettur['SO2'].mean()
avg_no2_mett=mettur['NO2'].mean()
avg_rspm_mett=mettur['RSPM/PM10'].mean()
print("Average SO2 in Mettur ",avg_so2_mett)
print("Average NO2 in Mettur",avg_no2_mett)
print("Average RSPM in Mettur",avg_rspm_mett)
```

```
Average SO2 in Mettur  8.429268292682927
Average NO2 in Mettur 23.185365853658535
Average RSPM in Mettur 52.72195121951219
```



```
#Average levels in Thoothukudi
import matplotlib.pyplot as plt
import seaborn as sns

thoo= df.loc[df['City/Town/Village/Area'] ==
'Thoothukudi']

avg_so2_thoo=thoo['SO2'].mean()
avg_no2_thoo=thoo['NO2'].mean()
avg_rspm_thoo=thoo['RSPM/PM10'].mean()
print("Average SO2 in Thoothukudi ",avg_so2_thoo)
print("Average NO2 in Thoothukudir",avg_no2_thoo)
print("Average RSPM in Thoothukudi",avg_rspm_thoo)
```

Output:

```
import matplotlib.pyplot as plt
import seaborn as sns
thoo= df.loc[df['City/Town/Village/Area'] == 'Thoothukudi']
avg_so2_thoo=thoo['SO2'].mean()
avg_no2_thoo=thoo['NO2'].mean()
avg_rspm_thoo=thoo['RSPM/PM10'].mean()
print("Average SO2 in Thoothukudi ",avg_so2_thoo)
print("Average NO2 in Thoothukudir",avg_no2_thoo)
print("Average RSPM in Thoothukudi",avg_rspm_thoo)
```

```
Average SO2 in Thoothukudi 12.988054607508532
Average NO2 in Thoothukudir 18.503412969283275
Average RSPM in Thoothukudi 83.45890410958904
```

Visualization:

#For SO2

```
import matplotlib.pyplot as plt
```

```
data=[avg_so2_tri,avg_so2_coi,avg_so2_mett,avg_so2_thoo,  
      avg_so2_chenn]
```

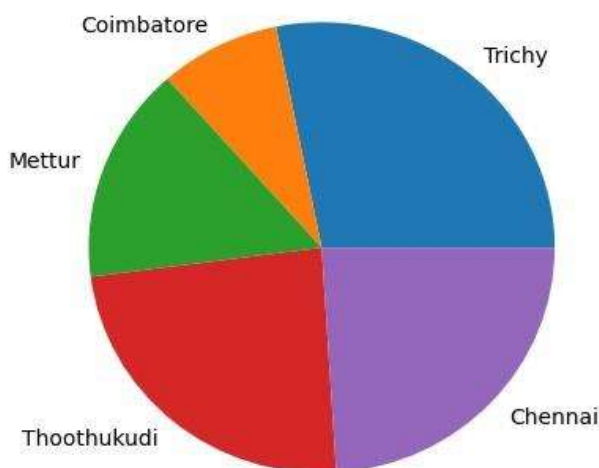
```
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi",  
         "Chennai"]
```

```
plt.pie(data,labels=mylabels)
```

```
plt.show()
```

Output:

```
import matplotlib.pyplot as plt  
data=[avg_so2_tri,avg_so2_coi,avg_so2_mett,avg_so2_thoo,avg_so2_chenn]  
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi","Chennai"]  
plt.pie(data,labels=mylabels)  
plt.show()
```





#For NO2

```
import matplotlib.pyplot as plt
```

```
data=[avg_no2_tri,avg_no2_coi,avg_no2_mett,avg_no2_thoo,  
      avg_no2_chenn]
```

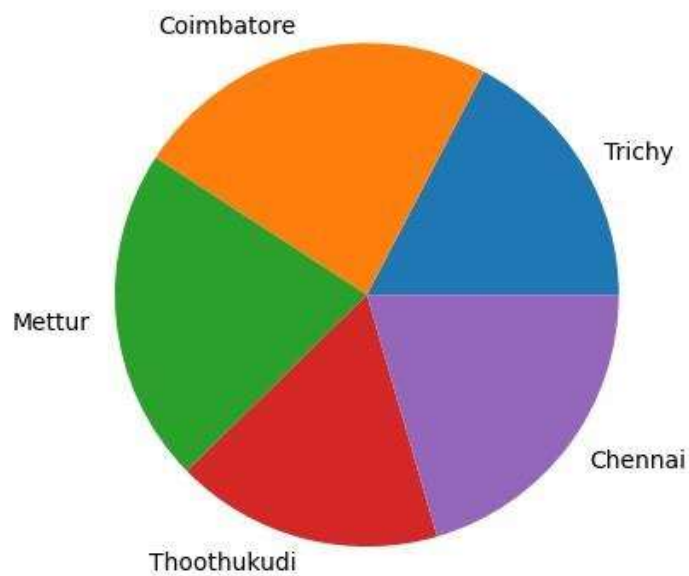
```
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi",  
         "Chennai"]
```

```
plt.pie(data,labels=mylabels)
```

```
plt.show()
```

Output:

```
import matplotlib.pyplot as plt  
data=[avg_no2_tri,avg_no2_coi,avg_no2_mett,avg_no2_thoo,avg_no2_chenn]  
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi","Chennai"]  
plt.pie(data,labels=mylabels)  
plt.show()
```



```
#For RSPM/PM10
```

```
import matplotlib.pyplot as plt
```

```
data=[avg_rspm_tri,avg_rspm_coi,avg_rspm_mett,avg_rspm_thoo,
```

```
      avg_rspm_chenn]
```

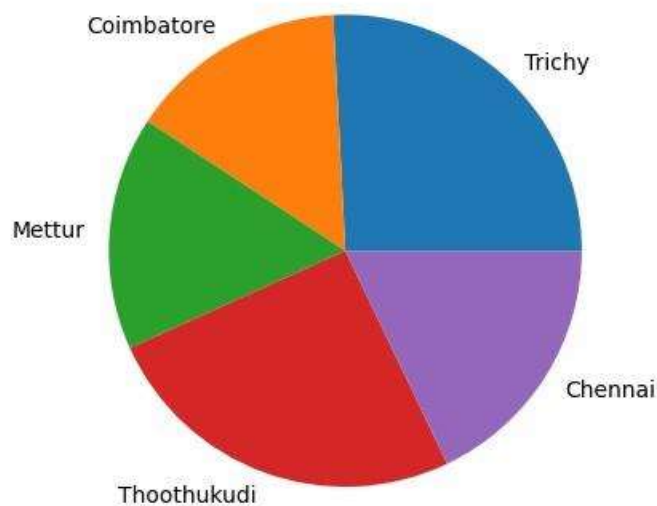
```
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi",  
          "Chennai"]
```

```
plt.pie(data,labels=mylabels)
```

```
plt.show()
```

Output:

```
import matplotlib.pyplot as plt  
data=[avg_rspm_tri,avg_rspm_coi,avg_rspm_mett,avg_rspm_thoo,avg_rspm_chenn]  
mylabels=["Trichy","Coimbatore","Mettur","Thoothukudi","Chennai"]  
plt.pie(data,labels=mylabels)  
plt.show()
```



## Conclusion:

=> By referring to the above pie charts the average RSPM/PM10 level of Trichy and Thoothukudi is nearly same and it is considered as High pollution areas,

=> Also mettur has Low level of RSPM level  
And it is considered as Low pollution area.