

AIR Q ASSESSMENT TN

Phase 4: Development Part 2

Dataset Link: <https://tn.data.gov.in/resource/location-wisedaily-ambient-air-quality-tamil-nadu-year-2014>

Loading Dataset:

Load your dataset into a pandas DataFrame, as we discussed in the previous response.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("C:\\Air Q TN.csv")
print(df['SO2'].mean())
```

Output:

```
11.503138075313808
```

```
print(df['NO2'].mean())
```

Output:

```
22.136775994417306
```

```
print(df['RSPM/PM10'].mean())
```

Output:

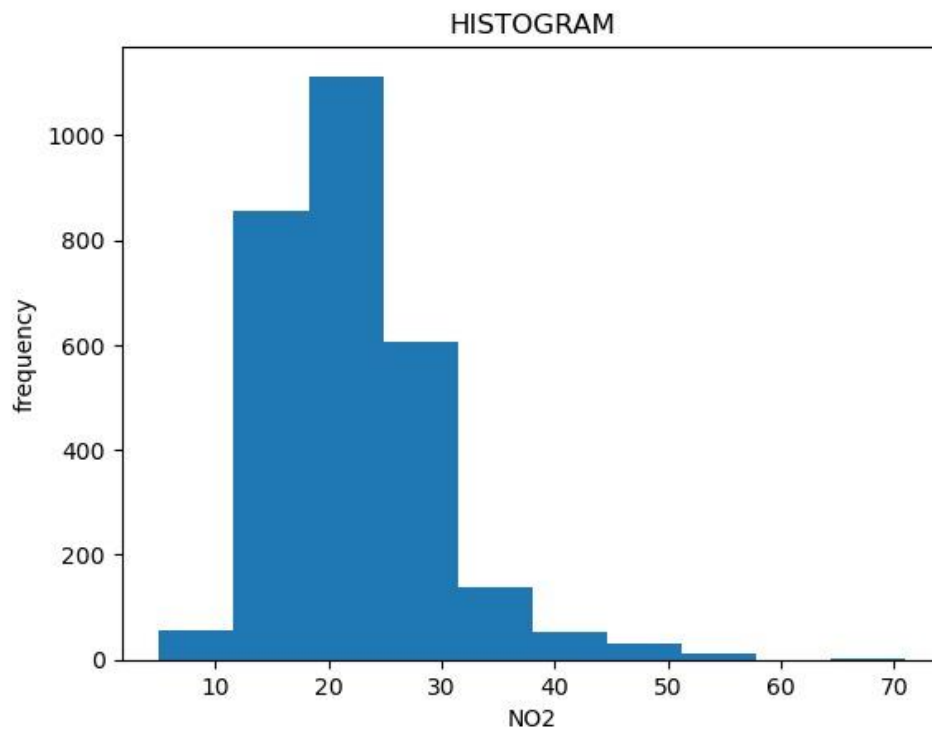
```
62.49426086956522
```

Histogram:

```
plt.title("HISTOGRAM")
```

```
plt.hist(df["NO2"])
plt.xlabel("NO2")
plt.ylabel("frequency")
plt.show()
```

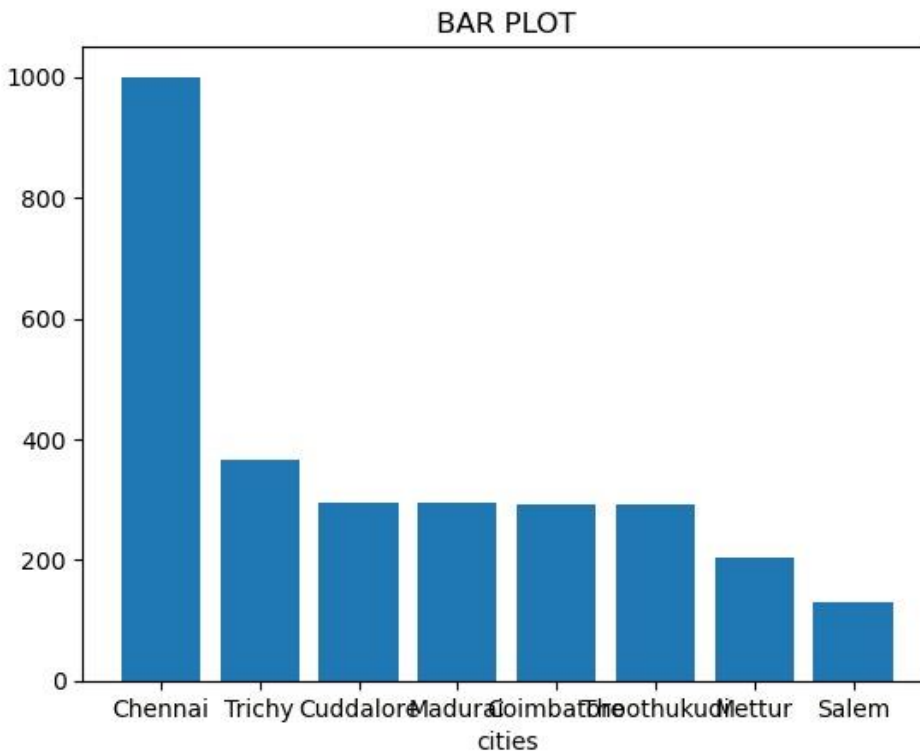
Output:



Bar Chart:

```
plt.title("BAR PLOT")
x=df["City/Town/Village/Area"].value_counts().nlargest(10)
plt.bar(x.keys(),x.values)
plt.xlabel("cities")
plt.ylabel("counts")
Plt.show()
```

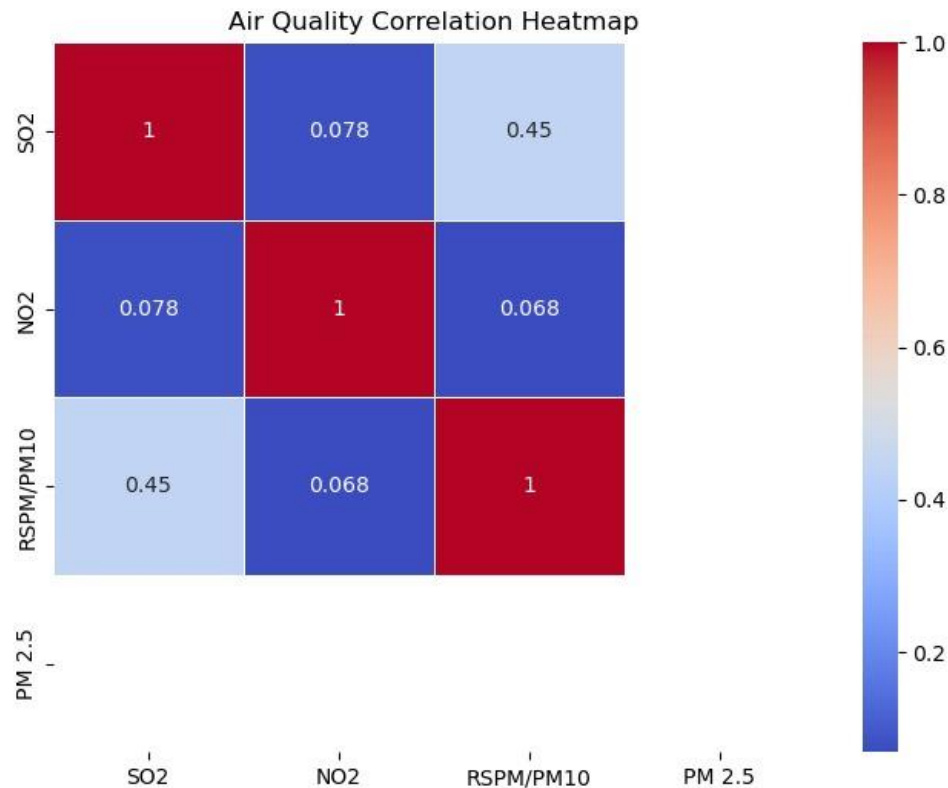
Output:



Heatmap:

```
numeric_columns = ['SO2', 'NO2', 'RSPM/PM10', 'PM 2.5']  
correlation_matrix = df[numeric_columns].corr()  
plt.figure(figsize=(8, 6))  
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm',  
linewidths=0.5)  
plt.title('Air Quality Correlation Heatmap')  
plt.show()
```

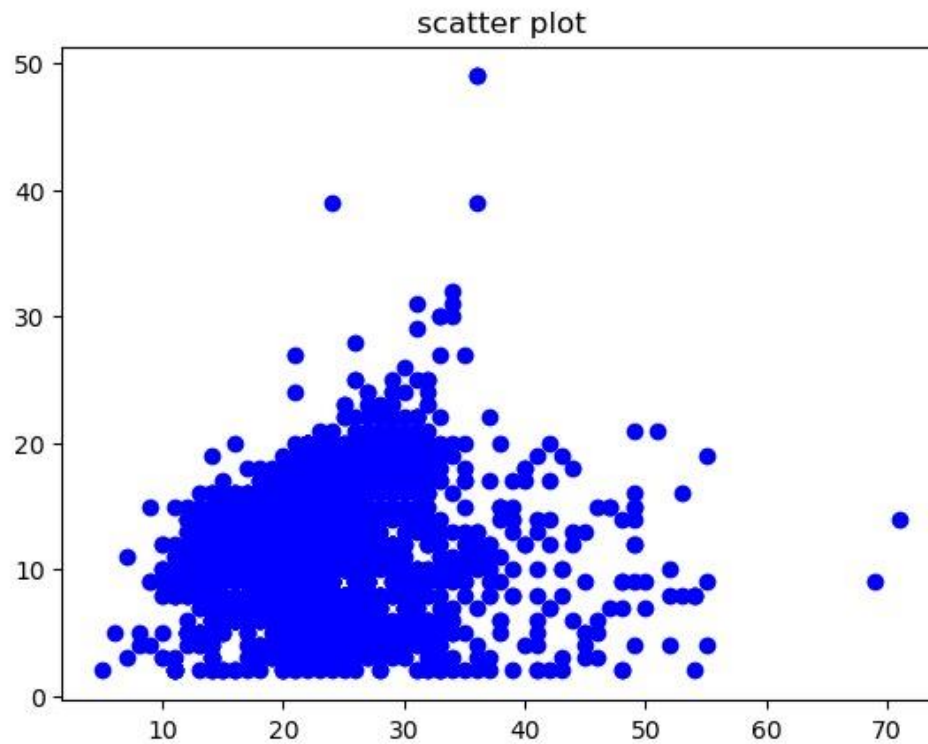
Output:



Scatter Plot:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("C:\\Air Q TN.csv")
plt.title("scatter plot")
plt.scatter(df["NO2"],df["SO2"],color='blue')
plt.show()
```

Output:



Line Plot:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("C:\\Air Q TN.csv")
x=df["NO2"]
y=x*2
plt.plot(x,y)
plt.show()
```

Output:

