Air Quality Analysis in Tamilnadu

Description:

- This script analyzes and visualizes air quality data for Tamil Nadu, India, using the `pandas` and `plotly` libraries. It loads the data from a CSV file, performs time series analysis, and creates a line plot for different air pollutants.

Importing Libraries:

The script begins by importing the necessary Python libraries for data analysis and visualization:

- `pandas`: A powerful data manipulation and analysis library.
- `plotly.express`: A high-level interface for creating a variety of complex plots.
- `plotly.io`: A library for configuring Plotly settings.
- `plotly.graph_objects`: A library for creating complex and sophisticated plots.

Setting Plotly Template:

- The default plotly template is set to "plotly_white" to provide a white background for the plots, enhancing their visibility.

Loading Data:

- The script loads air quality data from a CSV file located at "/kaggle/input/airqualityintamilnadu/cpcb_dly_aq_tamil_nadu-2014.csv" into a pandas Data Frame named `data`.

Data Processing:

- The 'Sampling Date' column in the Data Frame is converted to datetime format using the `pd.to_datetime()` function, facilitating time-based analysis.

Data Summary:

A summary of the data, including descriptive statistics, is printed using the `describe()` method to provide insights into the general characteristics of the dataset.

Creating Plotly Figure:

- An empty plotly figure (`fig`) is created, which will be populated with line plots representing the concentrations of different pollutants over time.

Plotting Time Series Data:

- Line plots are generated for each pollutant (SO2, NO2, and RSPM/PM10) using a loop to iterate through the list of pollutants.
- The 'Sampling Date' is set as the x-axis, and the corresponding pollutant concentration values are set as the y-axis for each plot.

Updating Layout and Labels:

- The plot's layout is updated with a title, x-axis label, and y-axis label for better interpretation and understanding of the visualization.

Displaying the Figure:

- The final plot is displayed using the `show()` function to visualize the time series data for different air pollutants.

Conclusion:

- This script serves as a basic framework for loading, processing, and visualizing air quality data, enabling users to gain insights into pollutant concentrations over time. Adjustments can be made to customize the visualization further or to accommodate different datasets and requirements.

Code :

```
import pandas as pd import
plotly.express as px import plotly.io
as pio import plotly.graph_objects as
go pio.templates.default =
"plotly_white"
 data =
pd.read_csv("/kaggle/input/airqualityintamilnadu/cpcb_dly_aq_tamil_nadu-
2014.csv")
print(data.head())
```

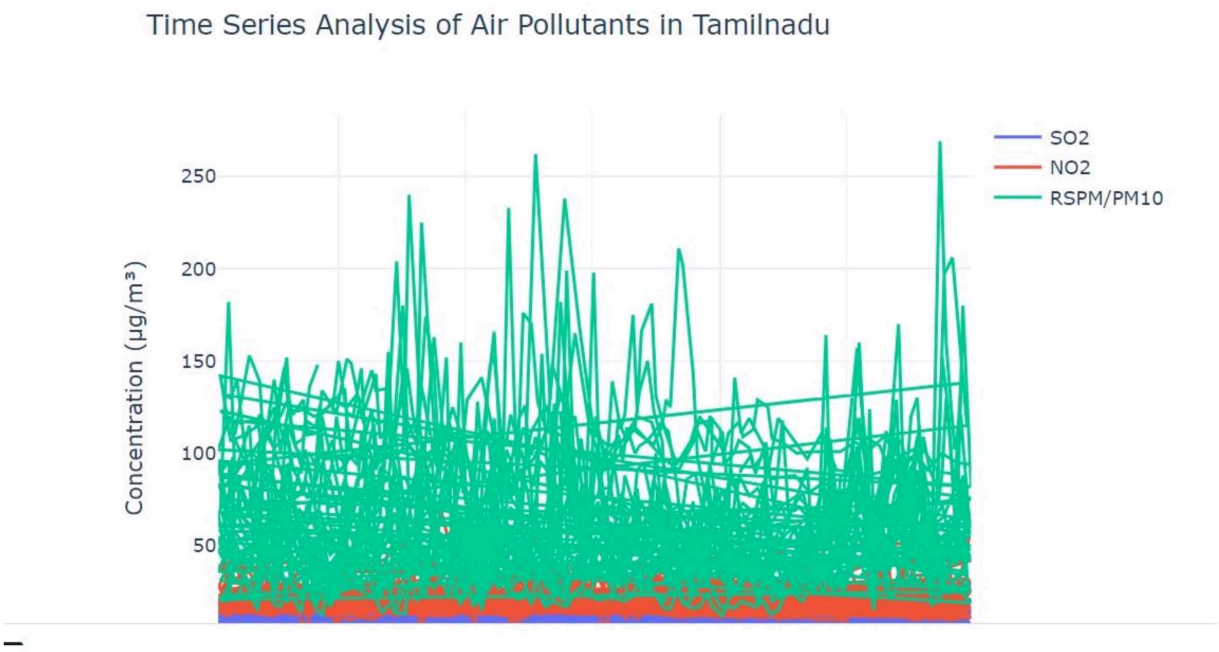
```
data['Sampling Date'] = pd.to_datetime(data['Sampling Date'])
print(data.describe())
```

```
fig = go.Figure()
 for pollutant in
['SO2','NO2','RSPM/PM10']:
    fig.add_trace(go.Scatter(x=data['Sampling Date'], y=data[pollutant],
mode='lines',                                name=pollutant))


fig.update_layout(title='Time Series Analysis of Air Pollutants in Tamilnadu',
xaxis_title='Date', yaxis_title='Concentration (µg/m³)') fig.show()
```

Output:

```
...      Stn Code Sampling Date       State City/Town/Village/Area  \
     0        38      01-02-14  Tamil Nadu               Chennai
     1        38      01-07-14  Tamil Nadu               Chennai
     2        38      21-01-14  Tamil Nadu               Chennai
     3        38      23-01-14  Tamil Nadu               Chennai
     4        38      28-01-14  Tamil Nadu               Chennai

                       Location of Monitoring Station  \
     0  Kathivakkam, Municipal Kalyana Mandapam, Chennai
     1  Kathivakkam, Municipal Kalyana Mandapam, Chennai
     2  Kathivakkam, Municipal Kalyana Mandapam, Chennai
     3  Kathivakkam, Municipal Kalyana Mandapam, Chennai
     4  Kathivakkam, Municipal Kalyana Mandapam, Chennai

                                     Agency Type of Location   SO2   NO2  \
     0  Tamilnadu State Pollution Control Board  Industrial Area  11.0  17.0
     1  Tamilnadu State Pollution Control Board  Industrial Area  13.0  17.0
     2  Tamilnadu State Pollution Control Board  Industrial Area  12.0  18.0
     3  Tamilnadu State Pollution Control Board  Industrial Area  15.0  16.0
     4  Tamilnadu State Pollution Control Board  Industrial Area  13.0  14.0

        RSPM/PM10  PM 2.5
     0       55.0     NaN
     1       45.0     NaN
     2       50.0     NaN
     3       46.0     NaN
     4       42.0     NaN
```

```
...            Stn Code                  Sampling Date          SO2          NO2  \
count   2879.000000                           2879  2868.000000  2866.000000
mean     475.750261  2014-06-28 11:39:14.567558144    11.503138    22.136776
min       38.000000            2014-01-02 00:00:00     2.000000     5.000000
25%      238.000000            2014-03-28 00:00:00     8.000000    17.000000
50%      366.000000            2014-06-26 00:00:00    12.000000    22.000000
75%      764.000000            2014-09-22 00:00:00    15.000000    25.000000
max      773.000000            2014-12-31 00:00:00    49.000000    71.000000
std      277.675577                            NaN     5.051702     7.128694

           RSPM/PM10  PM 2.5
count   2875.000000     0.0
mean      62.494261     NaN
min       12.000000     NaN
25%       41.000000     NaN
50%       55.000000     NaN
75%       78.000000     NaN
max      269.000000     NaN
std       31.368745     NaN
```



Conclusion:

- This script serves as a basic framework for loading, processing, and visualizing air quality data, enabling users to gain insights

into pollutant concentrations over time. Adjustments can be made to customize the visualization further or to accommodate different datasets and requirements.