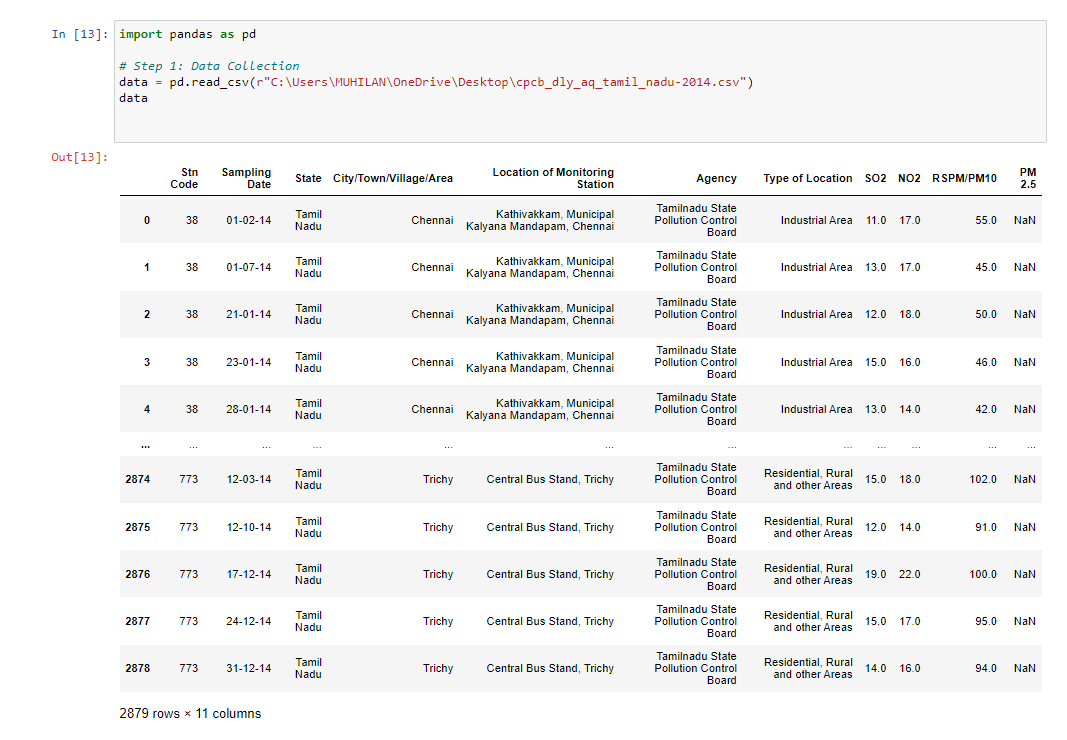
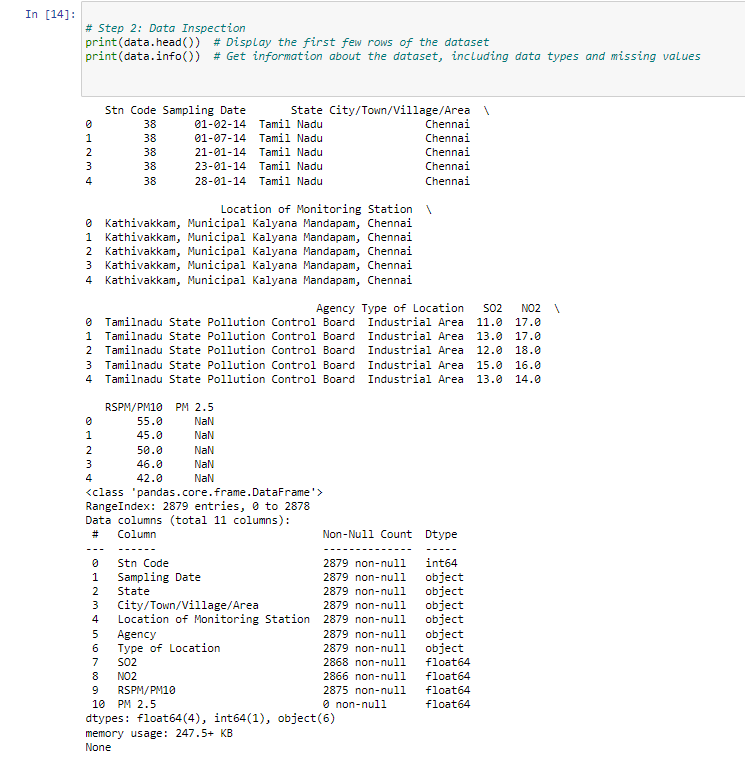
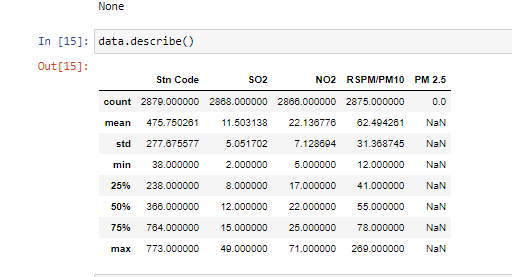
AIR QUALITY ANALYSIS IN TAMILNADU

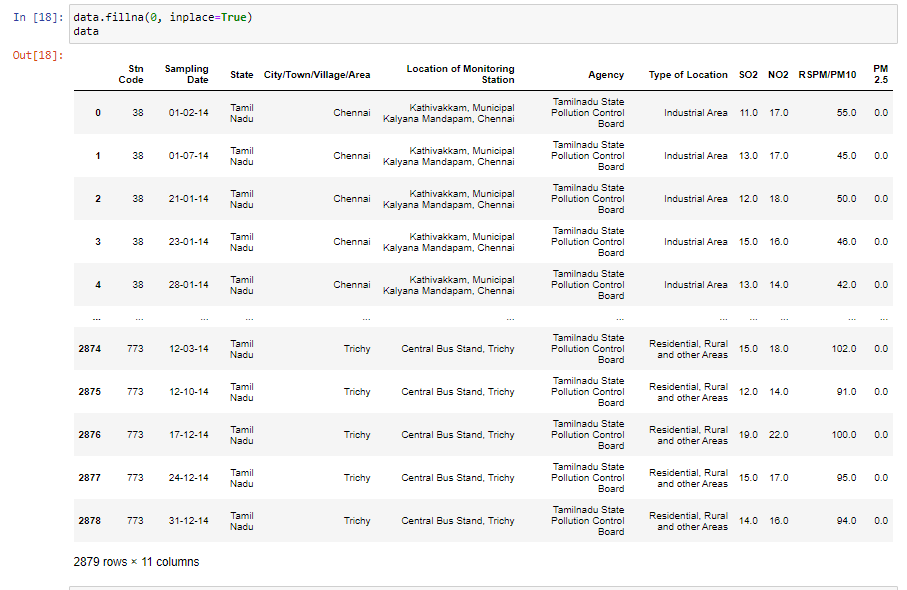
PHASE – 5

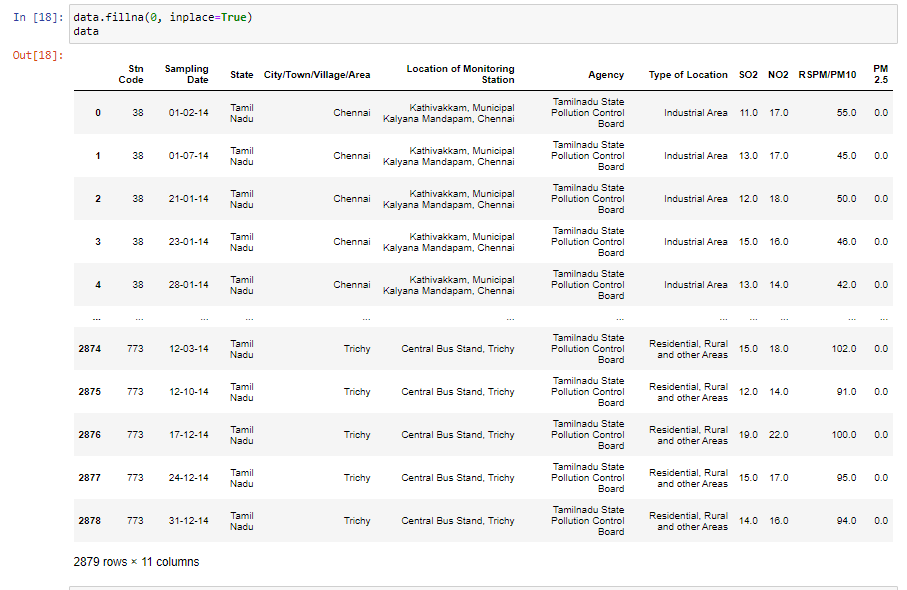
PREPROCESSING AND LOADING THE DATASET:











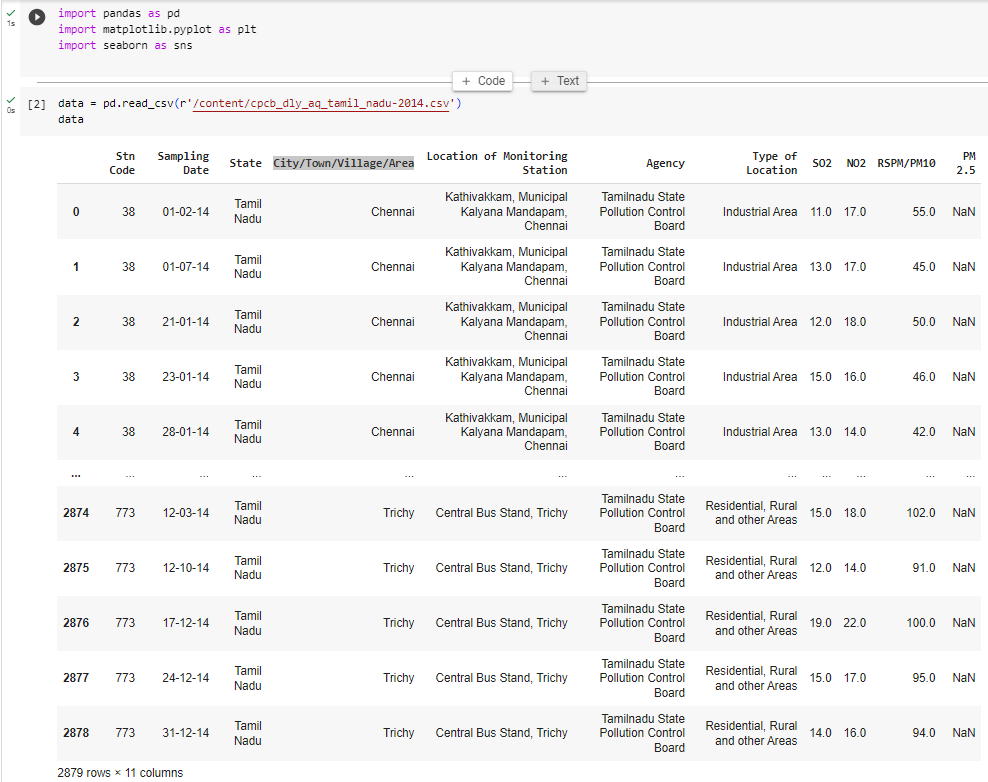
IMPORT LIBRARIES:

* you need to import the necessary libraries.

**LOAD AND PREPROCESS DATA**:

Load your air quality data into a pandas Data Frame and perform any necessary preprocessing. For this example, let's assume you have a CSV file named "/content/cpcb\_dly\_aq\_tamil\_nadu-2014.csv "

**OUTPUT:**



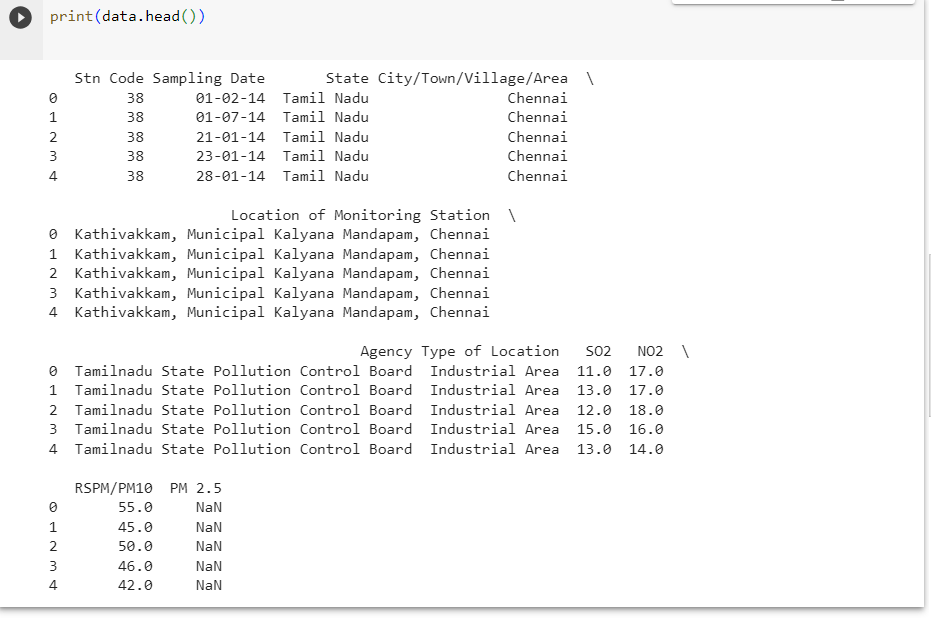
**CHECKING THE STRUCTURE OF THE DATA:**

By using the command, we can get the head of the given data set.

**COMMAND:**

print(data.head())

**OUTPUT:**



**CALCULATE AVERAGE POLLUTION** **LEVEL:**

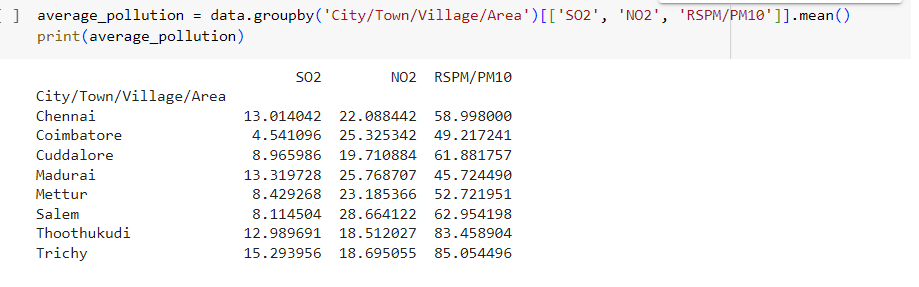
Calculate the average SO2, NO2, and RSPM/PM10 levels across different monitoring stations, cities, or areas.

COMMAND:

average\_pollution = data.groupby('city')[['SO2', 'NO2', 'RSPM/PM10']].mean()

print(average\_pollution)

OUTPUT:



**IDENTIFY POLLUTION TRENDS:**

You can identify pollution trends by visualizing the data. Here's an example of how to create some visualizations.

COMMAND:

sns.barplot(x='SO2', y='city', data=data)

plt.title('Average SO2 Levels by City')

plt.xlabel('SO2 Levels')

plt.ylabel('City')

plt.show()

sns.scatterplot(x='NO2', y='RSPM/PM10', data=data)

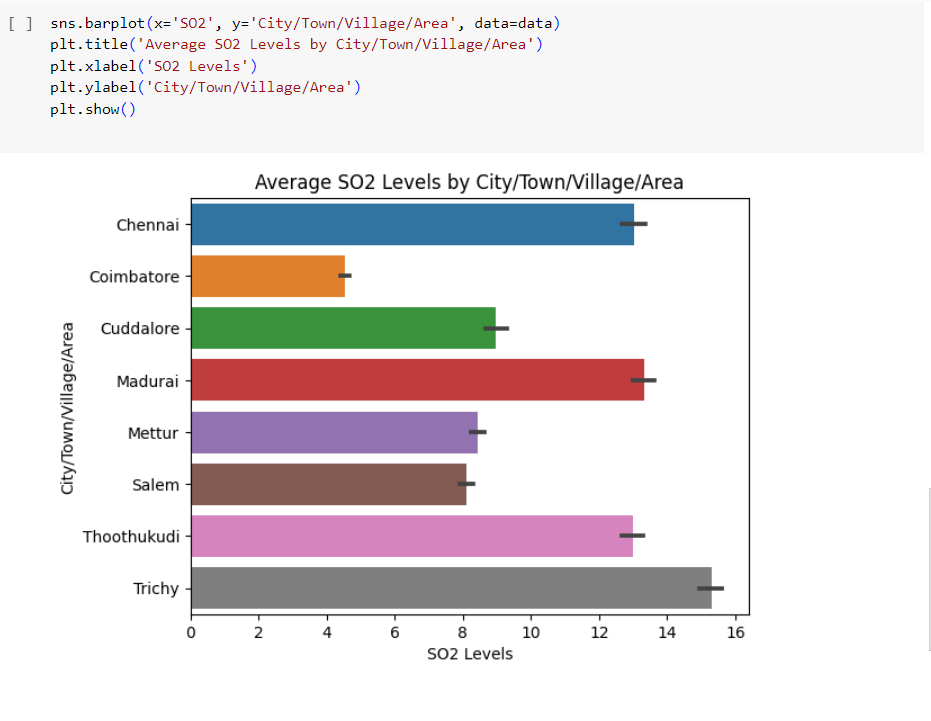
plt.title('Relationship between NO2 and RSPM/PM10 Levels')

plt.xlabel('NO2 Levels')

plt.ylabel('RSPM/PM10 Levels')

plt.show()

OUTPUT:



**IDENTIFY AREAS WITH HIGH POPULATION LEVELS:**

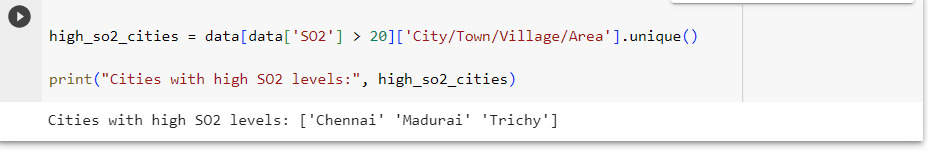
To identify areas with high pollution levels, you can filter the data based on specific criteria, such as cities with SO2 levels above a certain threshold.

COMMAND:

high\_so2\_cities = data[data['SO2'] > 20]['city'].unique()

print("Cities with high SO2 levels:", high\_so2\_cities)

OUTPUT:



SAVE AND DISPLAY VISUALIZATION:

You can save the visualizations as image files or display them directly in your Jupyter Notebook or script.

FURTHER ANALYSIS:

Depending on your specific goals, you can perform additional analyses and visualizations, such as time series analysis or geographical mapping of pollution levels.

Remember that the specific data and analysis may vary, so make sure to adapt the code to your dataset and analysis goals. Also, consider using other data visualization libraries or tools if they are more suitable for your needs.

INSIGHTS:

