Anna University Regional Campus Coimbatore

Anna University: Chennai-600 025

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

IBM Naan Mudhalvan Phase 3 Submission

Development Part-I

Title: AIR QUALITY ANALYSIS AND PREDICTION IN TAMILNADU

Name : PRAVEEN S

Register Number: 710021106027

Department :B.E.ECE

Sem/year :V/III

Air Quality Analysis and Prediction in Tamil Nadu

Objective:

The Objective of the project is to analyze and visualize the air quality data from the various monitoring stations in Tamil Nadu. The give dataset contains the measurements of the various gases that release into the atmosphere. Some of the gases that given in the dataset are Sulphur Dioxide(SO2), Nitrogen Dioxide(NO2) and Respirable particulate matter and these are measured in different cities, villages, towns. This Project aim is to gain the insight of air pollution trends, estimate the RSPM/PM10 levels based on SO2 and NO2 levels

Description of dataset:

The link for the chosen dataset is mentioned below,

https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014

The above dataset contains the combined version of air quality of Tamil Nadu. This contains the district wise dataset for the prediction of air quality parameter in the state of Tamil Nadu. This data was released by the Ministry of Environment and Forests and Central Pollution Control Board of India under the National Data Sharing and Accessibility Policy

1.Data Collection:

Monitoring Stations: Establish a network of air quality monitoring stations across Tamil Nadu. These stations should be strategically located in urban, industrial, and rural areas to capture a representative sample of air quality conditions.

- **Parameters:** Measure various air quality parameters, including particulate matter (PM 2.5 and PM 10), nitrogen dioxide(NO2), sulphur dioxide(SO2), carbon monoxide(CO), ozone(O3) and other volatile organic compounds.
- **Meteorological Data:** Collect meteorological data, such as temperature, humidity, wind speed,
  and wind direction, as these factors can influence air quality.
- **Historical Data:** Gather historical air quality data to establish trends and identify areas with chronic air quality problems.

2. Data Analysis:

Air Quality Index (AQI): Calculate the AQI for different locations in Tamil Nadu to provide a clear and understandable representation of air quality to the public.

- **Identify Hotspots:** Identify areas with consistently poor air quality, such as major cities or industrial zones, and pinpoint the key pollutants responsible.
- **Seasonal Trends:** Analyse seasonal variations in air quality, as well as the factors contributing to these variations, such as agricultural burning, weather conditions, or industrial activity.

3. Pollution Sources:

Industrial Emissions: Emissions from industrial facilities, such as factories and power plants, and assess compliance with emission standards. Examine

- **Vehicle Emissions:** Evaluate the impact of vehicular emissions on air quality, considering the prevalence of different types of vehicles and fuel types.
- **Agricultural Practices:** Investigate the role of agriculture in air quality, including the use of pesticides and burning of crop residues.
- **Waste Management:** Assess waste disposal practices and their impact on air quality, especially in urban areas.

4. Health Impact Assessment:

Collaborate with healthcare institutions to study the health effects of poor air quality on the population of Tamil Nadu. Identify vulnerable groups, such as children, the elderly, and individuals with pre-existing respiratory conditions, and assess their exposure and health outcomes.

5. Policy and Regulation:

Review existing air quality regulations and policies in Tamil Nadu to identify gaps or areas for improvement. Develop or update regulations to control emissions from various sources, and enforce strict compliance measures.

6. Public Awareness:

Launch public awareness campaigns to educate residents about the health risks associated with poor air quality and ways to protect themselves. Provide real-time air quality information through websites, apps, and public displays.

7. Mitigation Strategies:

Implement pollution control technologies in industries and encourage the use of cleaner fuels. Promote sustainable urban planning, public transportation, and green spaces to reduce vehicle emissions and enhance air quality. Encourage agricultural practices that minimize burning and promote sustainable waste management.

8. International Cooperation:

Collaborate with neighbouring states and countries to address transboundary air pollution issues, especially during cross-border events like crop burning. This air quality analysis is the first part of a comprehensive strategy to improve air quality in Tamil Nadu. It is essential to monitor progress over time and adjust strategies as needed to ensure cleaner air for the people and the environment.

PROGRAM:

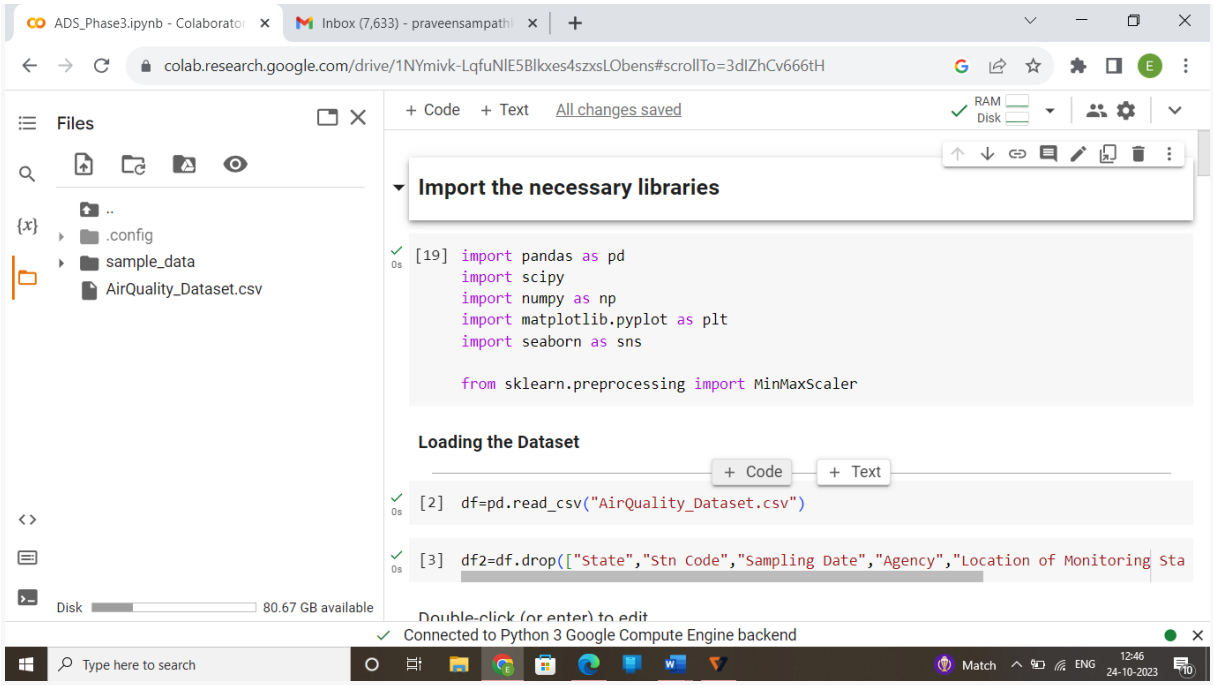
Import the necessary libraries:

```
import pandas as pd
import scipy
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import MinMaxScaler
```

Loading the Dataset:

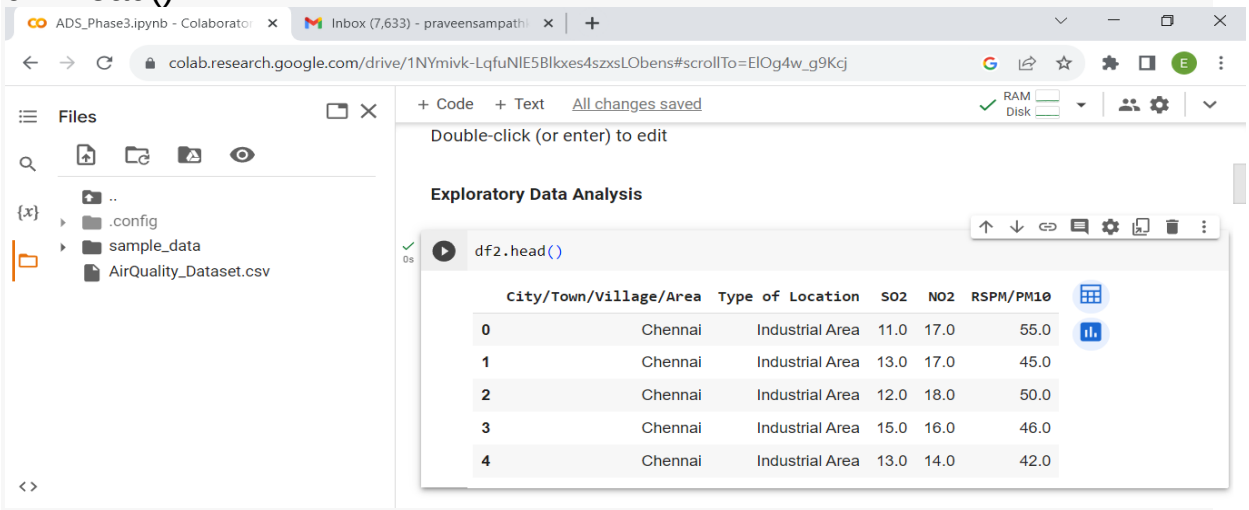
```
df=pd.read_csv("AirQuality_Dataset.csv")

df2=df.drop(["State","Stn Code","Sampling Date","Agency","Location of Monitoring Station","PM 2.5"], axis=1)
```



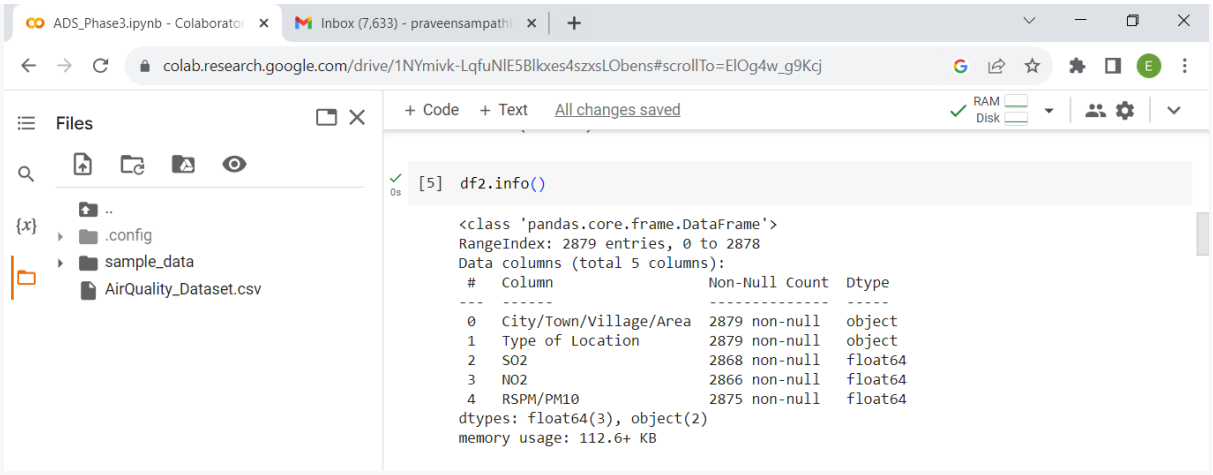
Exploratory Data Analysis:

df2.head()

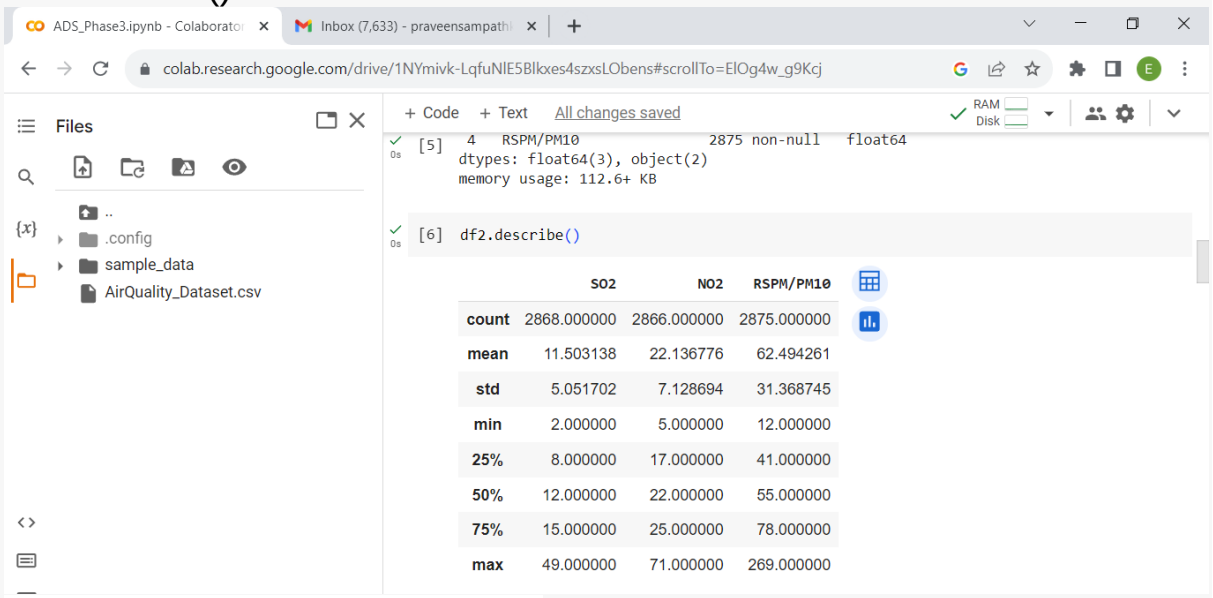


	City/Town/Village/Area	Type of Location	SO2	NO2	RSPM/PM10
0	Chennai	Industrial Area	11.0	17.0	55.0
1	Chennai	Industrial Area	13.0	17.0	45.0
2	Chennai	Industrial Area	12.0	18.0	50.0
3	Chennai	Industrial Area	15.0	16.0	46.0
4	Chennai	Industrial Area	13.0	14.0	42.0

df2.info()

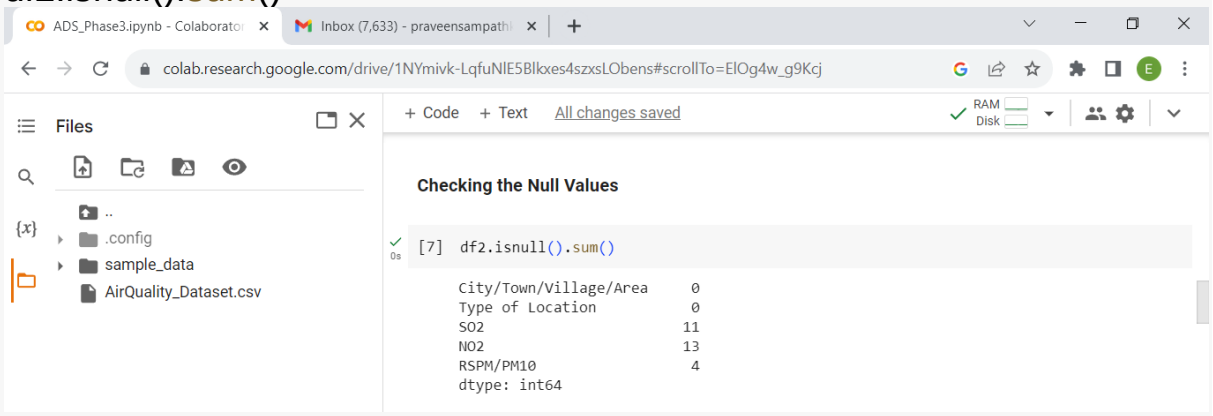


df2.describe()

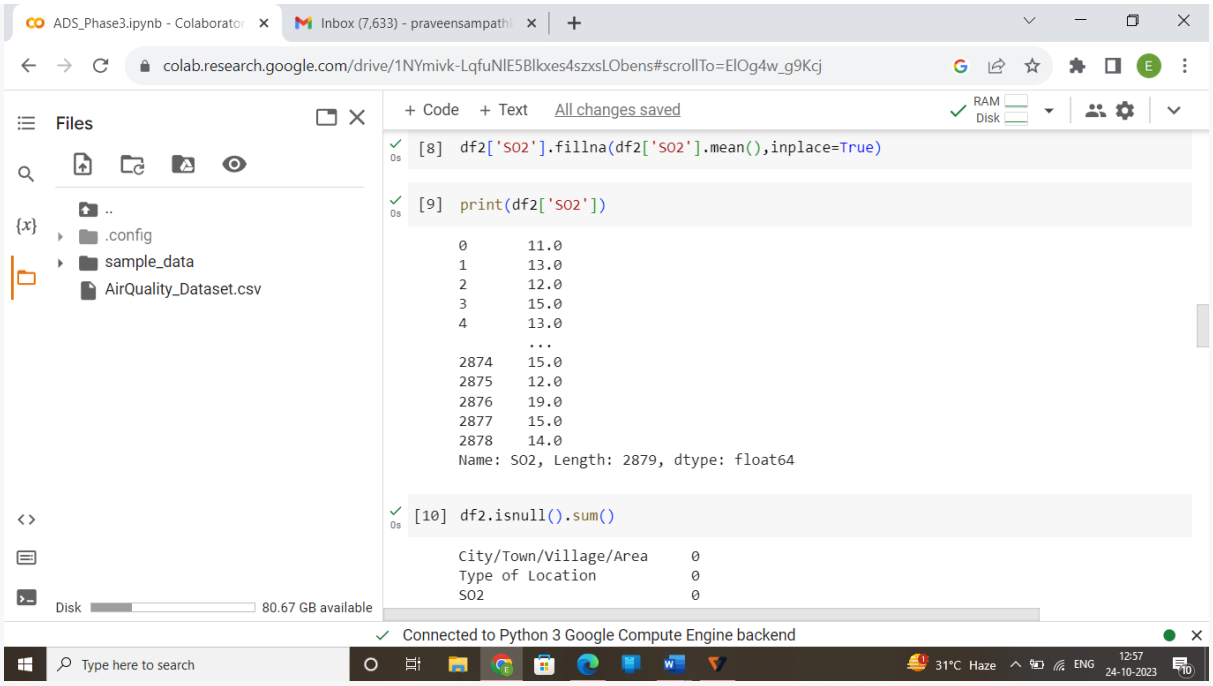


Checking the Null Values:

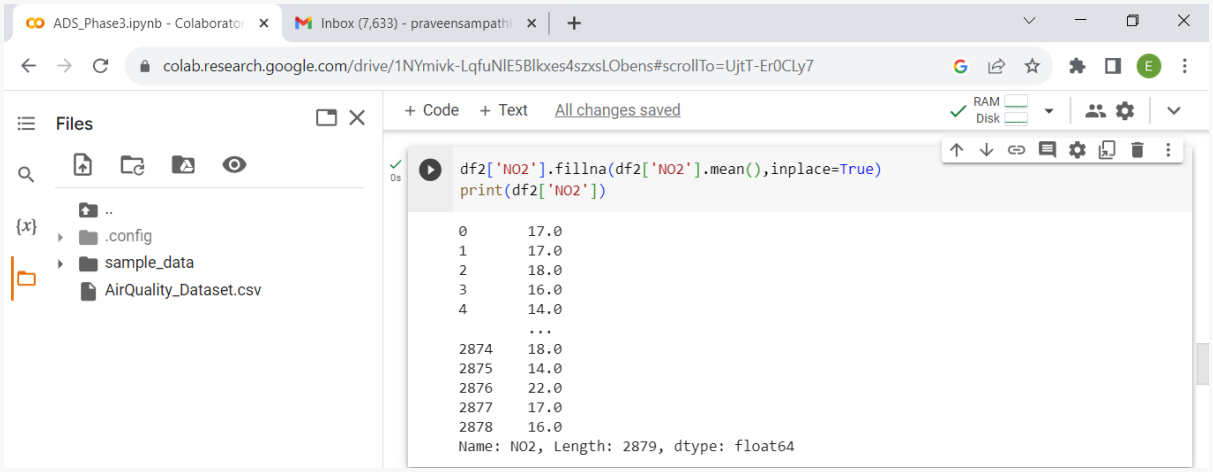
df2.isnull().sum()



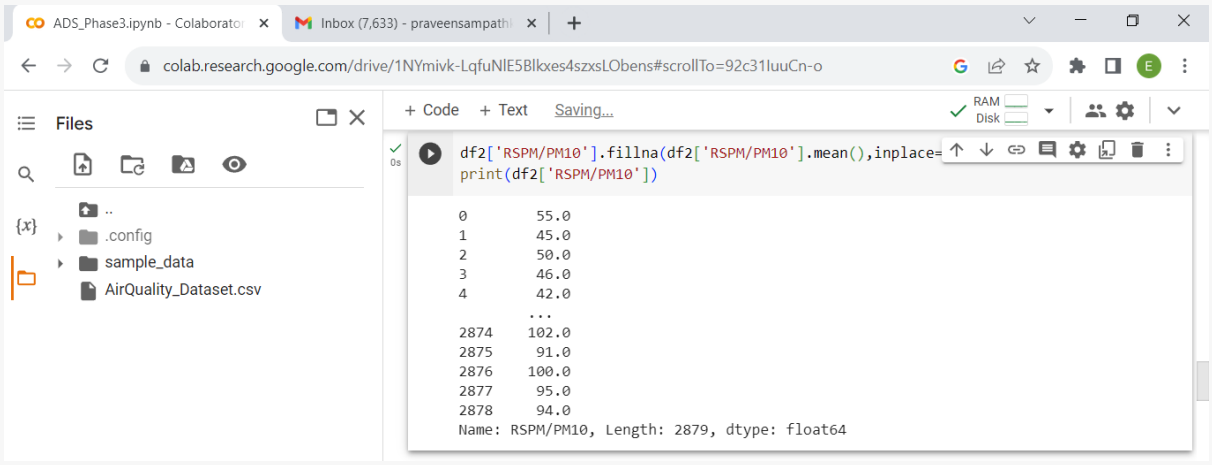
df2['SO2'].fillna(df2['SO2'].mean(),inplace=True)



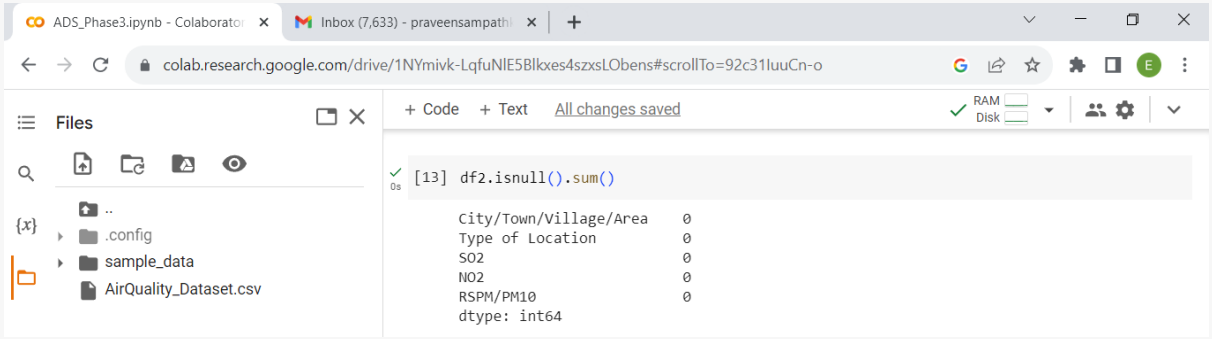
df2['NO2'].fillna(df2['NO2'].mean(),inplace=True)



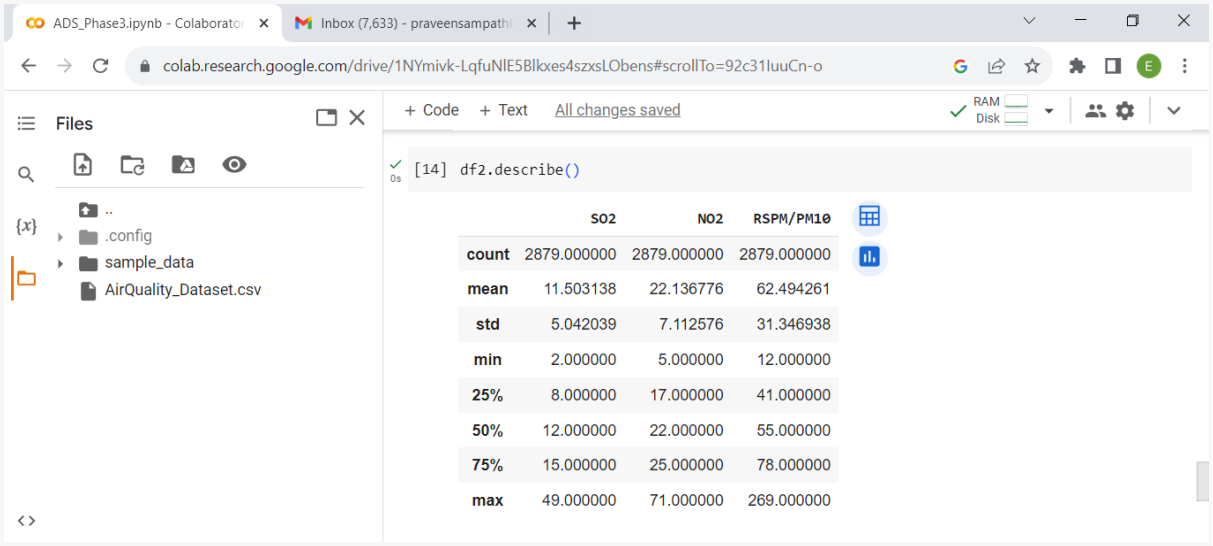
df2['RSPM/PM10'].fillna(df2['RSPM/PM10'].mean(),inplace=True)



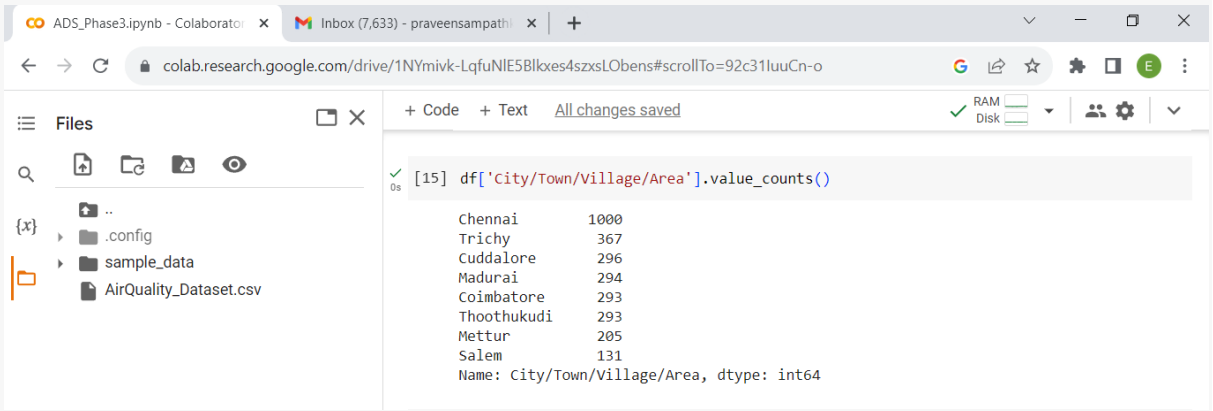
df2.isnull().sum()



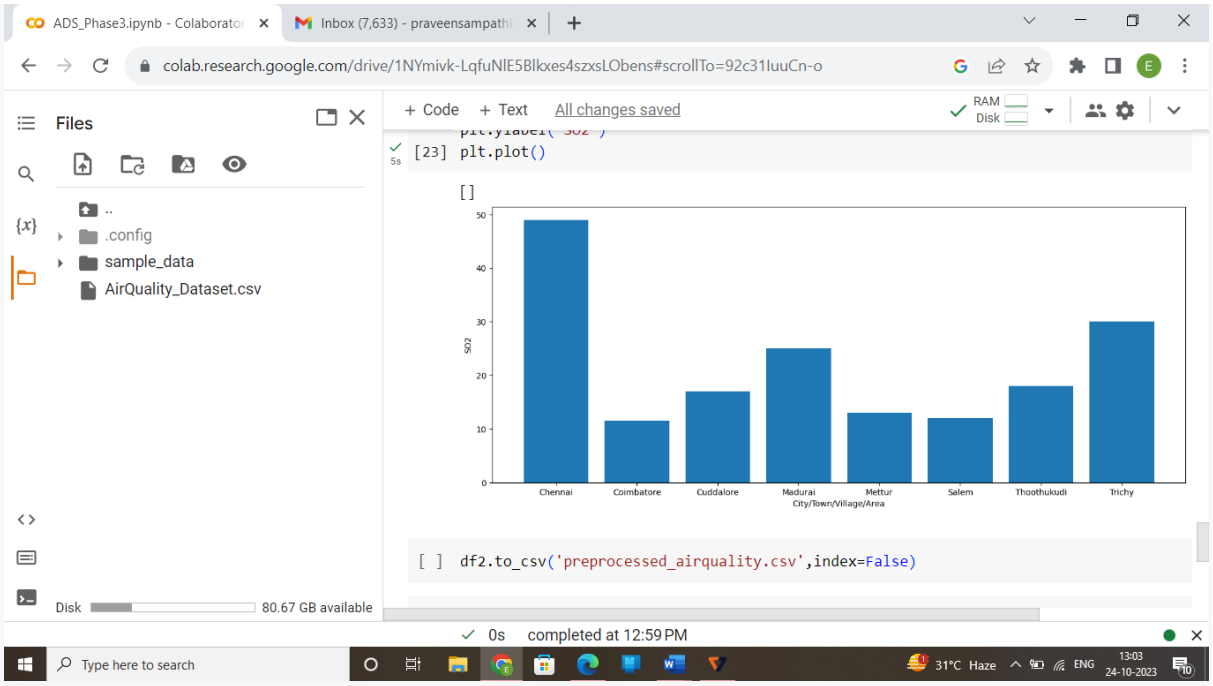
df2.describe()



df['City/Town/Village/Area'].value_counts()

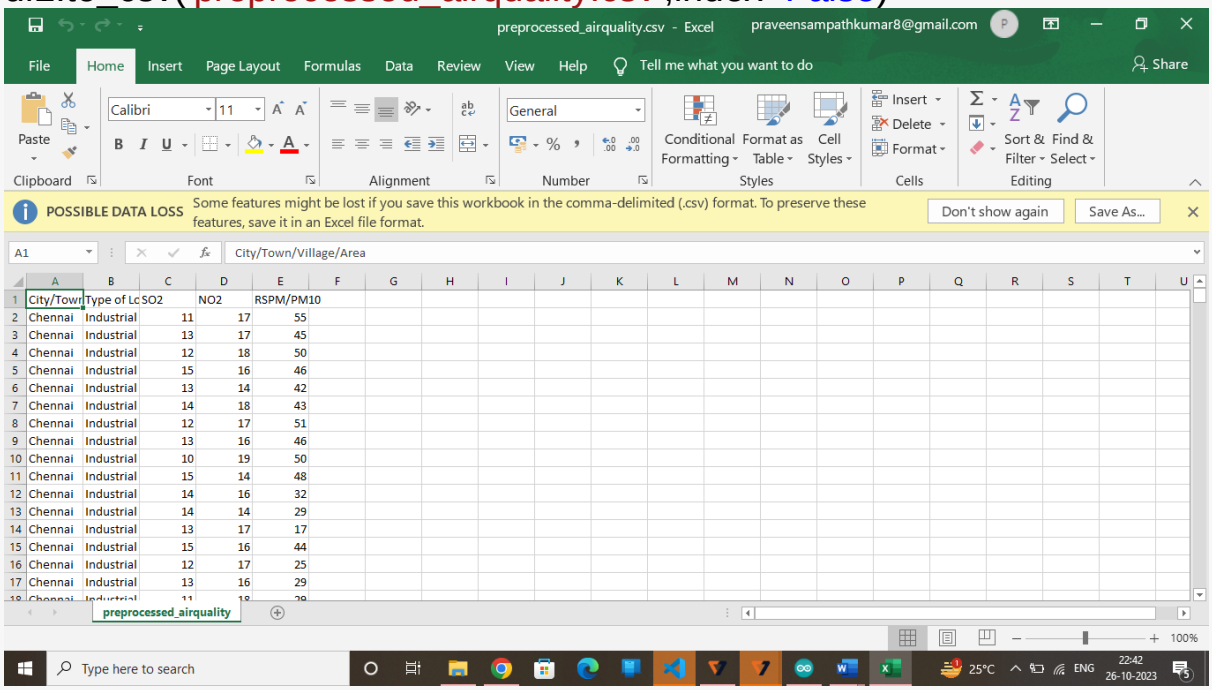



```
plt.figure(figsize=(15,6))
plt.bar(df2['City/Town/Village/Area'],df2['SO2'])
plt.xlabel('City/Town/Village/Area')
plt.ylabel('SO2')
plt.plot()
```



Saving the Pre-processed Data:

df2.to_csv('preprocessed_airquality.csv',index=False)



CONCLUSION:

In Conclusion, the data pre-processing is that it is a crucial and foundational step in the data analysis process. It involves tasks

such as data cleaning, data transformation and data reduction to ensure that the data is accurate, complete and suitable for analysis. Proper data pre-processing enhances the quality of the data and ultimately leads to more accurate and reliable results in data analysis and machine learning tasks. It is a critical phase that helps remove noise, handle missing values and prepare the data for modelling and proper decision-making.