Date

Air quality analysis

# Air quality analysis is the process of assessing the condition of the air in a specific area to determine the presence and concentration of various pollutants and contaminants. This analysis typically involves measuring parameters such as particulate matter (PM), gases like ozone (O3), nitrogen dioxide (NO2), sulfur dioxide (SO2), carbon monoxide (CO), and volatile organic compounds (VOCs). Monitoring stations and sensors are commonly used to collect data, and this information is crucial for assessing the health and environmental impacts of air pollution, as well as for regulatory purposes and public awareness. If you have specific questions or need more detailed information, feel free to ask!

# Code

# import pandas as pd

# # Load air quality data from a CSV file

# data = pd.read\_csv(‘air\_quality\_data.csv’)

# # Check the first few rows of the dataset

# print(data.head())

# # Calculate basic statistics

# mean\_concentration = data[‘PM2.5’].mean()

# max\_concentration = data[‘PM2.5’].max()

# min\_concentration = data[‘PM2.5’].min()

# # Filter data based on specific criteria

# high\_pollution\_days = data[data[‘PM2.5’] > 50]

# # Plotting air quality data (requires additional libraries like Matplotlib)

# import matplotlib.pyplot as plt

# data[‘Date’] = pd.to\_datetime(data[‘Date’])

# plt.plot(data[‘Date’], data[‘PM2.5’])

# plt.xlabel(‘Date’)

# plt.ylabel(‘PM2.5 Concentration’)

# plt.title(‘PM2.5 Concentration Over Time’)

# plt.show()

# Data Collection:

# Acquire air quality data from reliable sources or sensors. Common pollutants include PM2.5, PM10, NO2, CO, SO2, O3, and VOCs.

# Ensure data includes timestamps and geographical information (latitude and longitude).

# Data Preprocessing:

# Clean the data by handling missing values and outliers.

# Perform data transformations or aggregations if needed (e.g., hourly averages from minute-by-minute data).

# Descriptive Analysis:

# Calculate basic statistics: mean, median, variance, and standard deviation of pollutant concentrations.

# Identify temporal patterns and trends.

# Visualize data with time series plots, histograms, or box plots.

# Spatial Analysis:

# If you have geographic coordinates, create spatial plots or maps to visualize pollution patterns across different locations.

# Correlation Analysis:

# Assess relationships between different pollutants and meteorological variables like temperature and humidity.

# Use correlation coefficients and scatter plots to understand connections.

# Time Series Analysis:

# Apply time series analysis techniques such as decomposition, seasonal decomposition of time series (STL), and autoregressive integrated moving average (ARIMA) modeling to capture seasonality and trends.