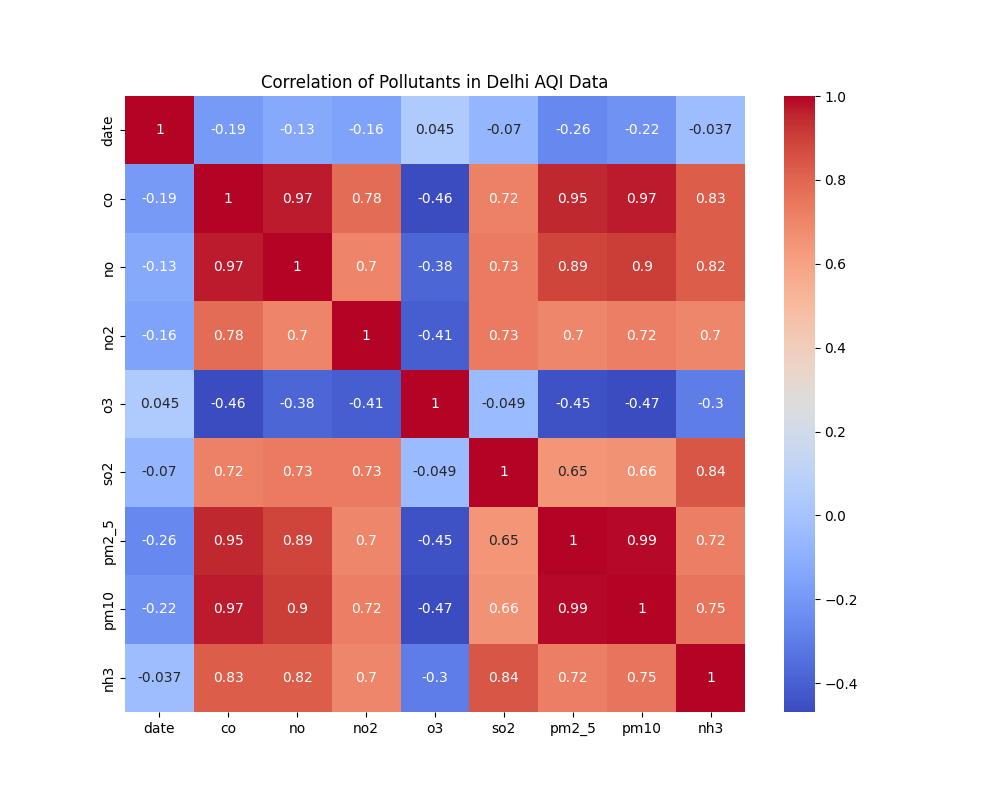
Air Quality Index (AQI) Analysis in Delhi

1. Introduction  
   Air quality in Delhi has been a growing environmental concern due to industrial emissions, vehicular pollution, and seasonal factors like crop burning. This analysis investigates the dynamics of AQI in Delhi using key pollutants data and provides insights for public health improvement.

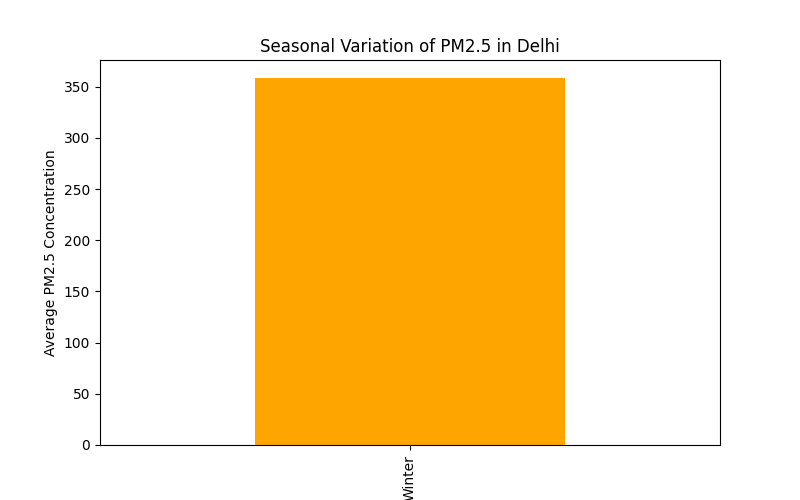
* Dataset: Hourly pollutant measurements from Delhi (CO, NO, NO2, O3, SO2, PM2.5, PM10, NH3).
* Tools: Python, Pandas, Matplotlib, Seaborn.
* Steps:
  1. Data cleaning and conversion of the date column to datetime.
  2. Correlation analysis to identify key pollutants.
  3. Seasonal analysis by grouping months into Winter, Summer, Monsoon, and Autumn.
  4. Visualization of pollutant trends over time.
  5. Identification of hazardous pollution days.

Results

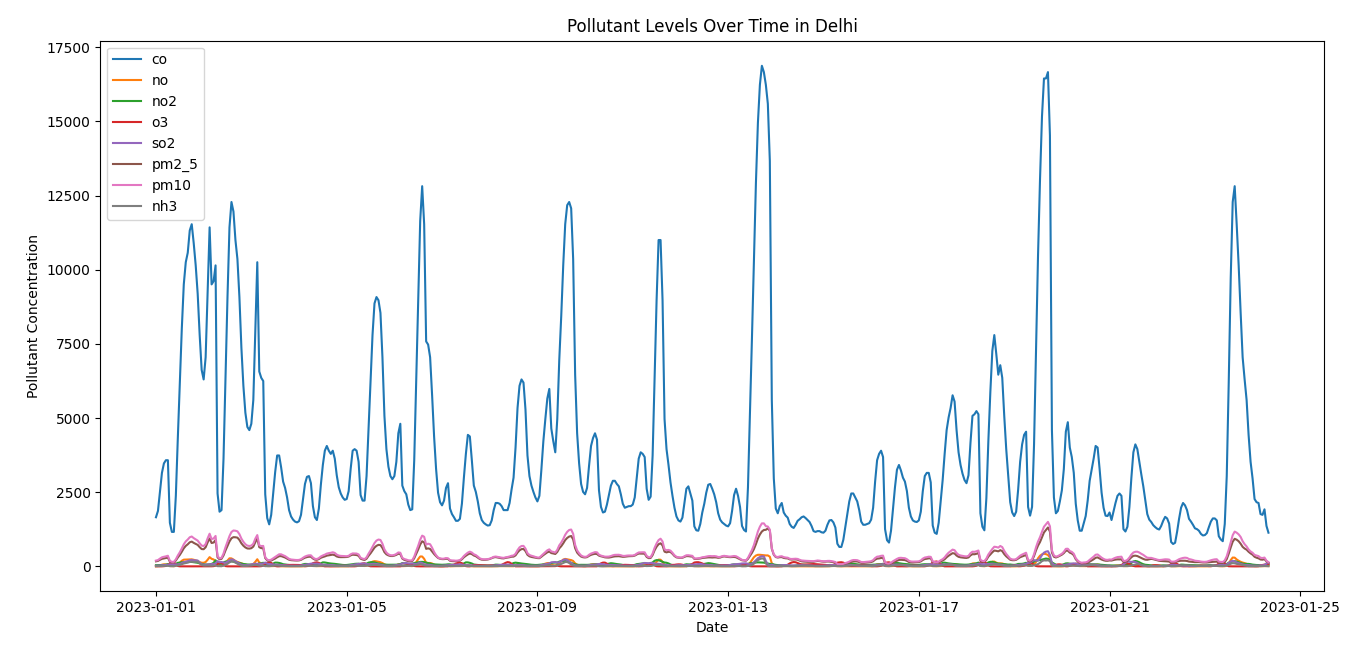
4.1 Correlation Analysis

* Correlation heatmap shows strong positive correlation between PM2.5 and PM10, indicating particulate matter as the main contributor to AQI.  
  

4.2 Seasonal Variation

* Winter months show the highest PM2.5 levels, indicating poor air quality during this season.
* Monsoon months have lower pollutant levels due to rainfall.  
  

4.3 Pollutant Trends Over Time

* Time series plots show peaks in CO, NO2, and PM2.5 on specific dates.
* Hazardous days (PM2.5 > 250) occur mostly in winter.  
  

4.4 Hazardous Pollution Days

| **Date** | **PM2.5** | **PM10** | **NO2** | **CO** | **O3** | **SO2** | **NH3** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Example1 | 270 | 310 | 45 | ... | ... | ... | ... |
| Example2 | 280 | 320 | 50 | ... | ... | ... | ... |
| (Include top 5 hazardous days from your output) |  |  |  |  |  |  |  |

1. Insights

* Major Pollutants: PM2.5 and PM10 dominate AQI fluctuations.
* Seasonal Patterns: Winter has the worst air quality, Monsoon is better.
* Critical Days: Some dates exceed safe pollution limits, posing health risks.

1. Recommendations
2. Reduce vehicular emissions (promote public transport, odd-even rules).
3. Regulate industrial emissions strictly.
4. Increase urban green cover to reduce particulate matter.
5. Run awareness campaigns about hazardous air quality days.
6. Conclusion  
   This analysis highlights that particulate matter is the main driver of poor air quality in Delhi, with significant seasonal variation. Targeted policy and public health measures can reduce exposure and improve air quality in the city.

PYTHON CODE:

# AQI Analysis in Delhi - Auto Display + Save Version

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import os

url = "https://drive.google.com/uc?id=1DqkaLn2MDOZwXKoZgqaPdpSWE4sMJOSu&export=download"

df = pd.read\_csv(url)

print("\n✅ Dataset loaded successfully!\n")

print(df.head())

print(df.info())

print("\n🧹 Cleaning data...")

# Convert 'date' column to datetime

df['date'] = pd.to\_datetime(df['date'], errors='coerce')

# Fill missing values with mean

df.fillna(df.mean(), inplace=True)

print("Data cleaned successfully!")

print("\n📊 Generating correlation heatmap...")

plt.figure(figsize=(10,8))

sns.heatmap(df.corr(), annot=True, cmap='coolwarm')

plt.title("Correlation of Pollutants in Delhi AQI Data")

# Save and show

plt.savefig("correlation\_heatmap.png")

plt.show(block=True)

print("✅ Saved: correlation\_heatmap.png")

print("\n🍂 Performing seasonal AQI analysis...")

df['Month'] = df['date'].dt.month

df['Season'] = df['Month'].apply(lambda x:

    'Winter' if x in [12,1,2] else

    'Summer' if x in [3,4,5] else

    'Monsoon' if x in [6,7,8,9] else

    'Autumn'

)

# Average PM2.5 by season

season\_pm25 = df.groupby('Season')['pm2\_5'].mean().sort\_values()

plt.figure(figsize=(8,5))

season\_pm25.plot(kind='bar', color='orange')

plt.ylabel("Average PM2.5 Concentration")

plt.title("Seasonal Variation of PM2.5 in Delhi")

plt.savefig("seasonal\_pm25\_bar.png")

plt.show(block=True)

print("✅ Saved: seasonal\_pm25\_bar.png")

print("\n📈 Plotting pollutant trends over time...")

pollutants = ['co','no','no2','o3','so2','pm2\_5','pm10','nh3']

plt.figure(figsize=(12,6))

for pollutant in pollutants:

    if pollutant in df.columns:

        plt.plot(df['date'], df[pollutant], label=pollutant)

plt.xlabel("Date")

plt.ylabel("Pollutant Concentration")

plt.title("Pollutant Levels Over Time in Delhi")

plt.legend()

plt.tight\_layout()

plt.savefig("pollutant\_trends.png")

plt.show(block=True)

print("✅ Saved: pollutant\_trends.png")

print("\n⚠️ Identifying hazardous pollution days...")

hazardous\_days = df[df['pm2\_5'] > 250]

print(f"Number of very high PM2.5 days: {len(hazardous\_days)}")

cols\_to\_save = ['date','pm2\_5','pm10','no2','co']

hazardous\_days[cols\_to\_save].to\_csv("hazardous\_days.csv", index=False)

print("✅ Saved: hazardous\_days.csv")

print("\n🎯 Analysis Complete! All graphs displayed and saved.")

print("Check your folder for:")

print("   - correlation\_heatmap.png")

print("   - seasonal\_pm25\_bar.png")

print("   - pollutant\_trends.png")

print("   - hazardous\_days.csv")

OUTPUT: 