Global Air Quality Data EDA

Problem Statement:

Air pollution is one of the most pressing global environmental challenges, directly impacting human health, ecosystems, and climate. Monitoring air quality helps policymakers and environmental agencies make informed decisions to control emissions and protect public health. In this project, we will explore global air quality datasets to identify pollution trends, seasonal variations, and city/country comparisons.

Dataset Link:

https://www.kaggle.com/datasets/rohanrao/air-quality-data-in-india

EDA Questions:

1. Which cities have the highest average AQI (Air Quality Index)?
2. How does AQI vary by month or season?
3. What pollutants are recorded, and what are their average levels?
4. Which city has the cleanest air on average?
5. Are there any missing data patterns in the dataset?
6. Which cities consistently exceed safe pollution limits?
7. How do pollution levels vary between weekdays and weekends?

Visualization Questions:

1. Bar chart of Top 10 most polluted cities by AQI.

2 Line plot showing monthly AQI trends for selected cities.

3 Heatmap showing correlation between pollutants (PM2.5, PM10, NO2, SO2, O3, CO).

4 Seasonal comparison plot of average AQI levels.

5 Stacked bar chart of pollutant contributions in different cities.

6 Map visualization showing geographic distribution of AQI values.

7 Boxplot comparing AQI across major cities.