

Experiment 2  
Hardik Garg  
2021300036  
BE Comps A - BATCH H

**Aim:** Analyze socio-economic dataset and show various analyses using advanced graphs and charts like boxPlot, word cloud etc.

### 1. Columns in the Dataset:

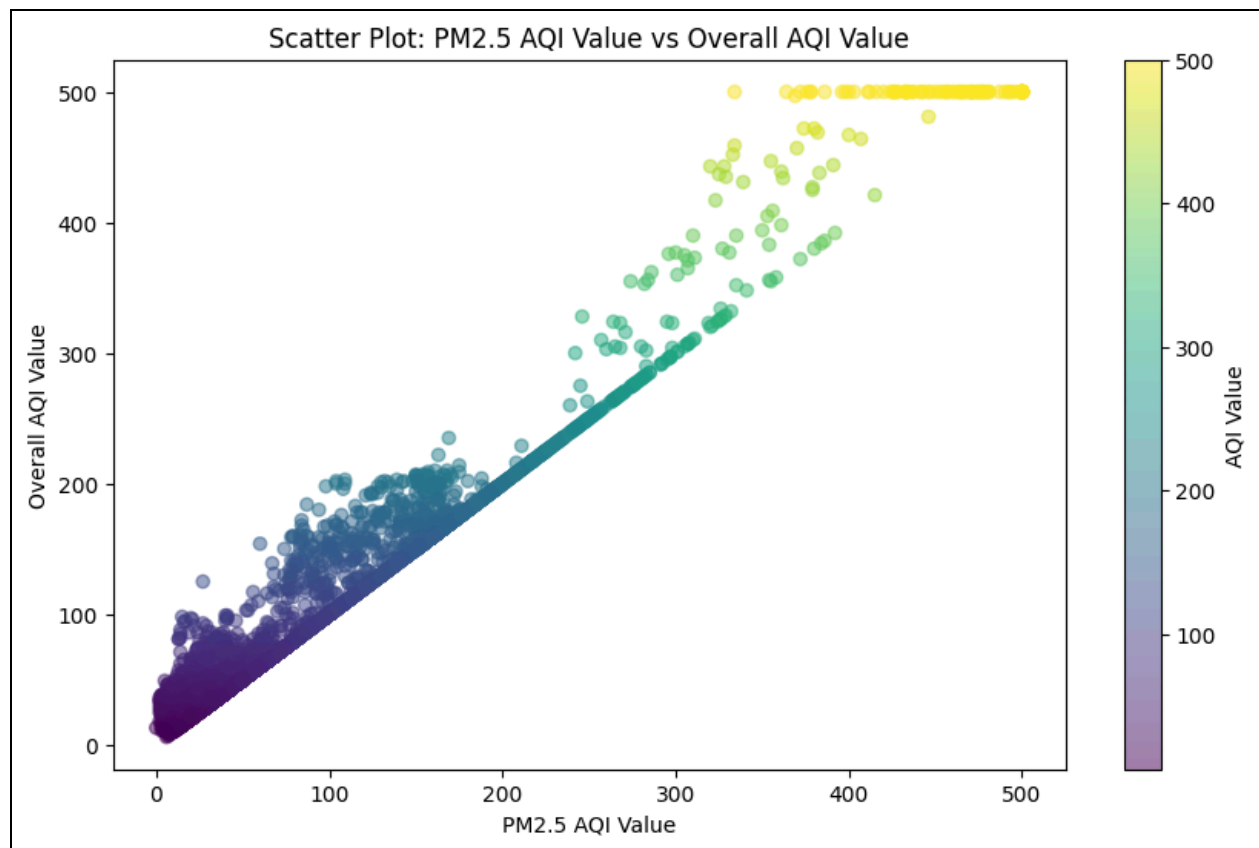
- **Country:** Represents the country where air quality data was recorded (object type).
- **City:** The specific city for which the air pollution data is available (object type).
- **AQI Value:** The overall Air Quality Index, indicating the level of air pollution (int64 type).
- **AQI Category:** The category of air quality (e.g., "Good," "Moderate") based on the AQI Value (object type).
- **CO AQI Value:** The AQI value specific to Carbon Monoxide (CO) (int64 type).
- **CO AQI Category:** The air quality category based on CO AQI value (object type).
- **Ozone AQI Value:** The AQI value specific to Ozone pollution (int64 type).
- **Ozone AQI Category:** The air quality category based on Ozone AQI value (object type).
- **NO2 AQI Value:** The AQI value specific to Nitrogen Dioxide (NO2) (int64 type).
- **NO2 AQI Category:** The air quality category based on NO2 AQI value (object type).
- **PM2.5 AQI Value:** The AQI value for fine particulate matter (PM2.5) (int64 type).
- **PM2.5 AQI Category:** The air quality category based on PM2.5 AQI value (object type).

### 2. Analysis Goals:

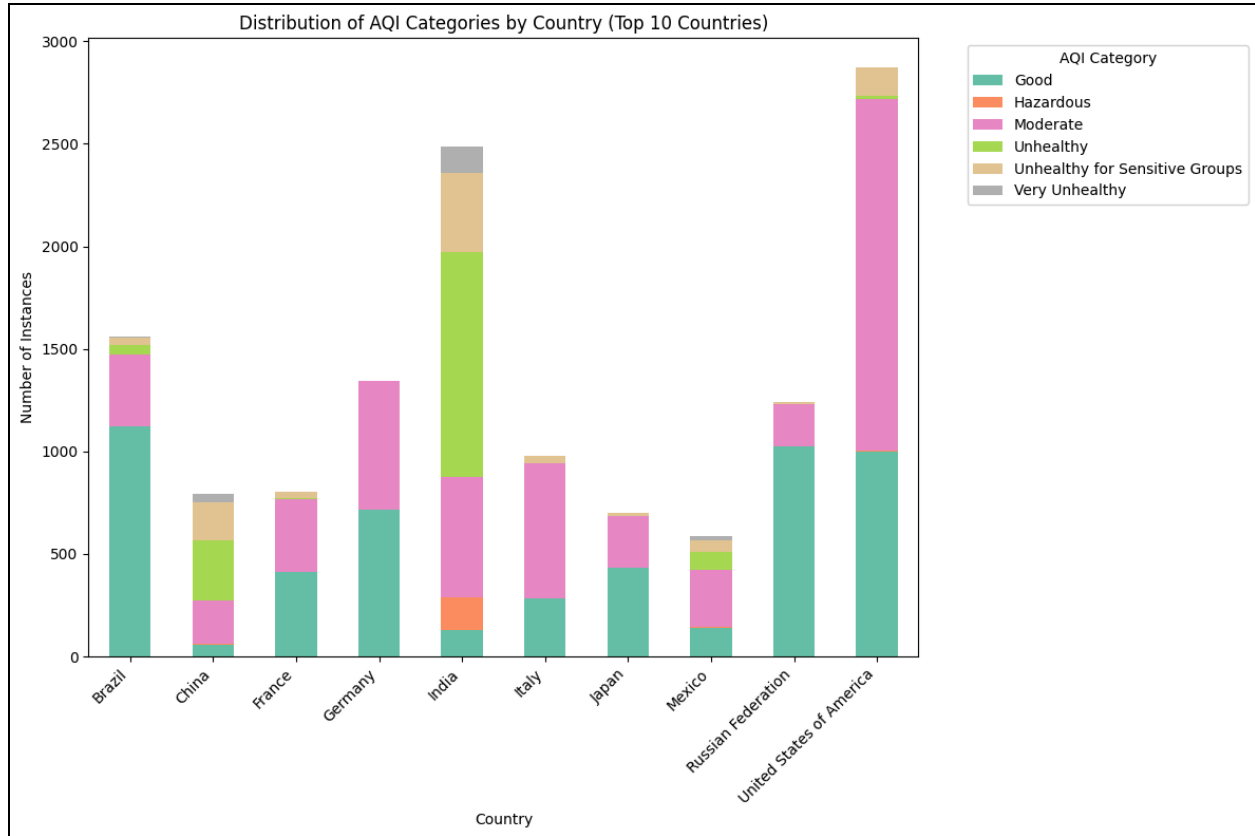
- **Country-Wise Pollution Analysis:** Group data by Country to observe regional air quality trends and identify countries with severe pollution.
- **Pollutant-Specific Analysis:** Compare AQI values for CO, Ozone, NO2, and PM2.5 across cities or countries to determine which pollutants contribute most to poor air quality.

- [illegible]

1. United States, India, Brazil and Mexico have been mentioned the most times as shown by this chart.



1. PM2.5 AQI Value is almost linearly related to the overall AQI value.



## Conclusion:

- **Primary Market Focus:** The dataset is heavily skewed towards the UK market, which should be the primary focus for any targeted marketing or sales strategies.
- **Volatile Sales Patterns:** The line chart highlights the importance of understanding the reasons behind the fluctuating sales quantities, which could involve deeper analysis of seasonal trends, product launches, or economic conditions during the periods in question.