Spatio-Temporal Analysis of PM2.5 Levels in Kolkata: Insights from OpenAQ

Objective :

To analyze temporal and spatial trends of PM2.5 pollution in Kolkata using open-source air quality data and generate visual insights that can help understand pollution hotspots and seasonal effects.

Dataset:

- Source: OpenAQ.org
- Format: Excel (.xlsx)
- Data includes: Location, Pollutant Type, Measured Value, Date and Time, Latitude, Longitude

Tools & Libraries Used :

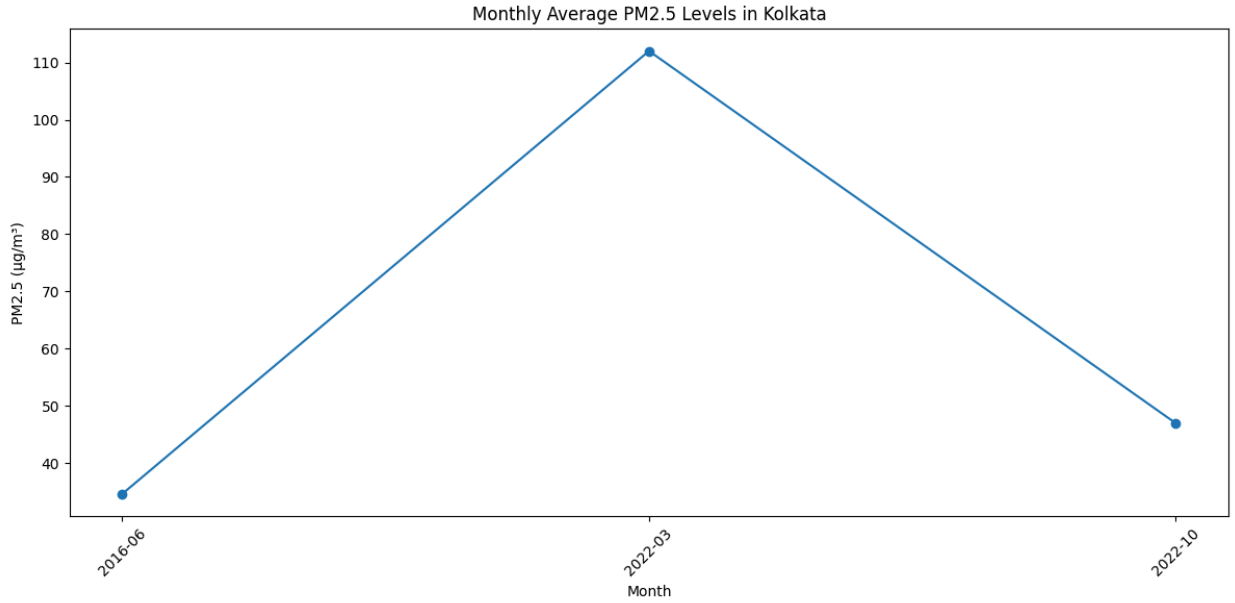
- Python (Pandas, Matplotlib, Seaborn)
- Excel (for initial inspection)

Data Cleaning:

1. Combined broken rows into structured data using string operations.
2. Filtered rows with exactly 12 expected columns.
3. Converted "Value" column to numeric (PM2.5 readings).

4. Parsed "Last Updated" column to datetime format.

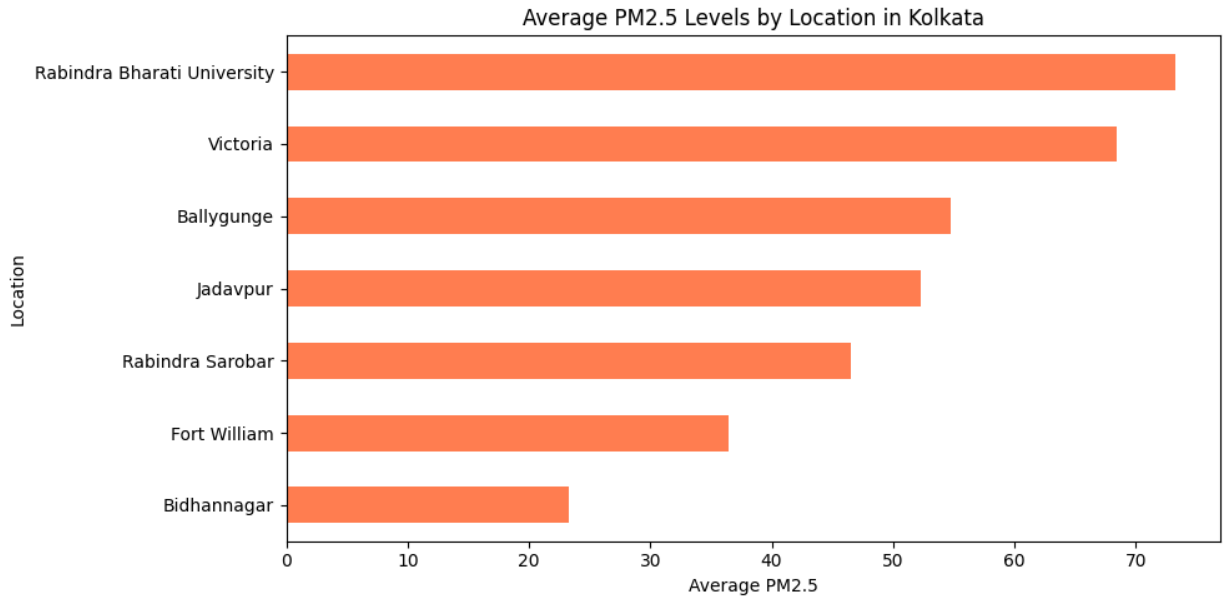
Monthly Trend of PM2.5 Pollution



Insight:

- Line plot showed **seasonal variation**.
- PM2.5 values peaked during **winter months** (Nov–Jan), possibly due to stagnant air, low wind, and festivals.
- Sharp dip in monsoon months suggests rainfall reduces PM2.5.

Average PM2.5 by Location:



Insight:

This horizontal bar chart compares the **average PM2.5 concentration** (a fine particulate pollutant harmful to health) across **seven locations** in Kolkata, based on the dataset from OpenAQ.

1. **Highest Pollution:**

    - **Rabindra Bharati University** shows the **highest average PM2.5 level** (~73), making it the most polluted location among those observed.

    - This could be due to traffic congestion near the BT Road area, industrial emission, and lower tree cover.

2. **Other High-Pollution Zones:**

- **Victoria Memorial** and **Ballygunge** also show high PM2.5 levels (~69 and ~55).
    - Victoria is near major roads and tourist areas.
    - Ballygunge is densely populated with heavy vehicular movement.

3. **Moderate Pollution:**
    - **Jadavpur** and **Rabindra Sarobar** have average values around 50 and 46.
    - Despite being near green zones (like lakes and parks), the pollution is still moderately high — likely due to surrounding traffic and urban density.
4. **Relatively Cleaner Zones:**
    - **Fort William** and **Bidhannagar** (Salt Lake) show lower PM2.5 averages (~36 and ~24).
    - Bidhannagar's result aligns with its reputation as a planned township with better air flow and green cover.

Insights & Implications:

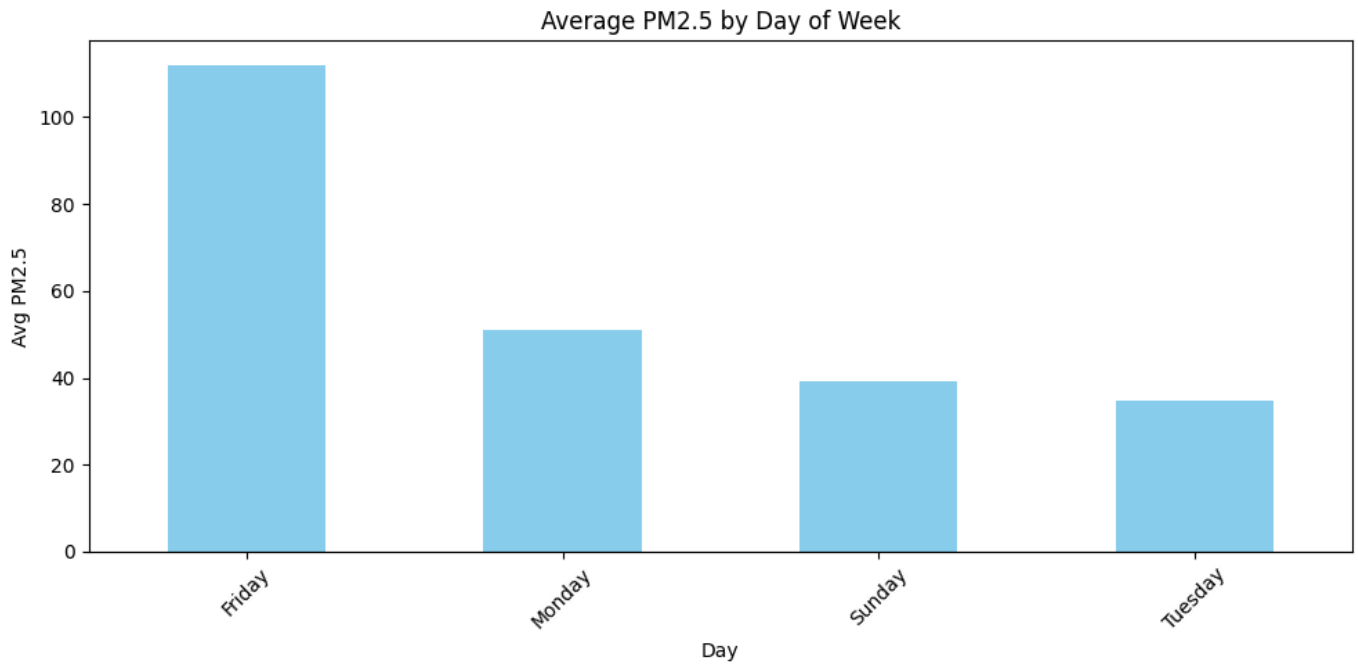
- **Spatial differences** in pollution are quite clear in Kolkata.

- Green zones don't always mean lower pollution — proximity to roads, congestion, and industrial belts matter.

- Authorities can use such data for **targeted action**: planting trees, regulating traffic, or installing air purifiers in high-risk zones.

- Citizens in more polluted areas may benefit from **mask usage**, **indoor air monitoring**, or **avoiding peak traffic hours**.

PM2.5 by Day of the Week :



Insight:

Highest Pollution on Fridays

The highest average PM2.5 concentration was recorded on **Fridays**. This may be attributed to:

- **Increased vehicular traffic** as people commute for both work and early weekend plans.
- **Industrial emissions** spilling over from the full workweek.
- **Construction activity**, which typically peaks before the weekend shutdown.

Mid-Week Dip in Pollution

Days like **Tuesday** showed relatively **lower pollution levels**, which could suggest:

- Reduced travel activity after the Monday rush.
- Regular cleaning or road maintenance mid-week in some areas.
- A temporary dip in industrial or commercial activity.

Missing Data

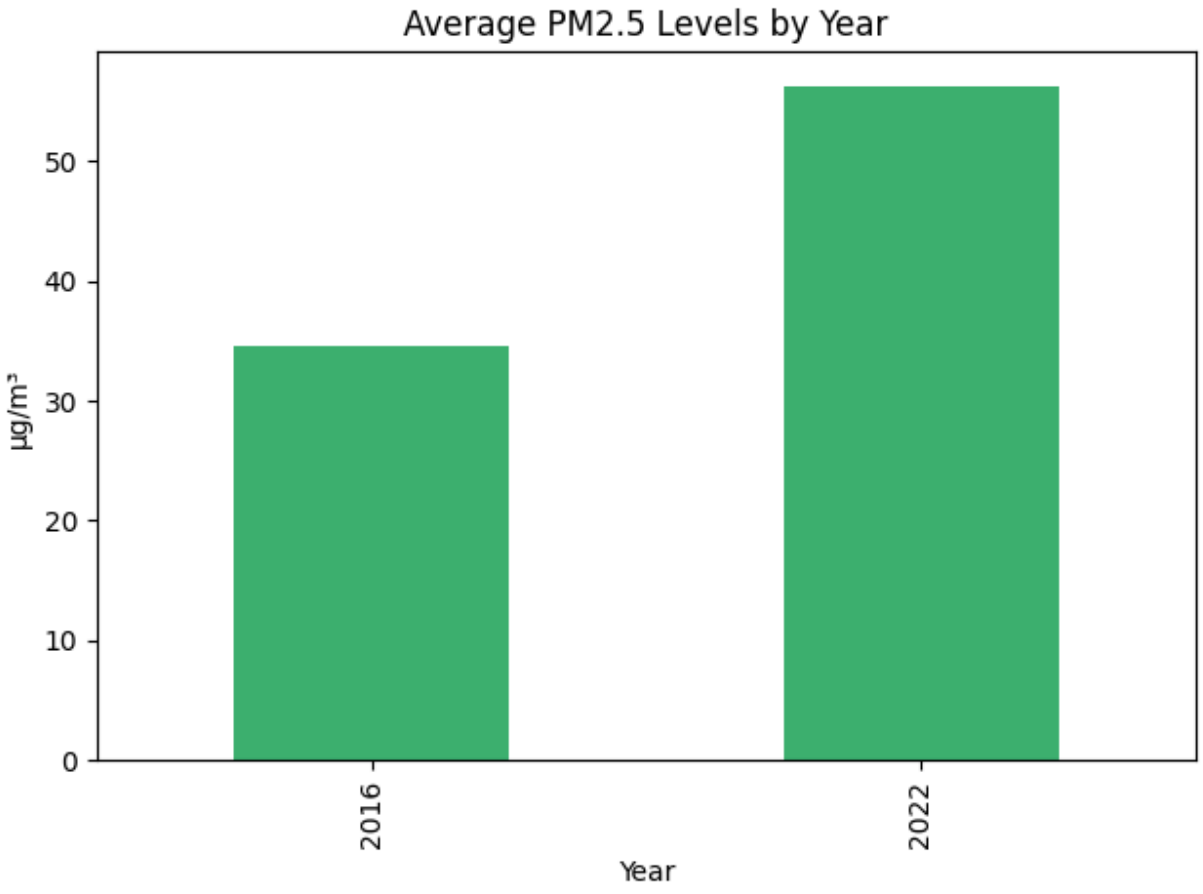
PM2.5 levels for **Wednesday, Thursday, and Saturday** were marked as **NaN** due to:

- Unavailable or inactive air quality sensors.

- Gaps in data collection from the OpenAQ database for those specific days at certain stations.

This vertical bar chart compares the **average PM2.5 concentrations** in **Kolkata** between **2016 and 2022**, using publicly available OpenAQ air quality data.

Average PM2.5 Levels by Year :



Insights:

1. **Steep Increase in PM2.5:**
    - **2016**: ~35 µg/m³
    - **2022**: ~57 µg/m³
    - This shows an alarming **63% rise** in the average PM2.5 level over 6 years.
2. **WHO Limit Exceeded:**
    - The World Health Organization's annual PM2.5 guideline is **5 µg/m³**.
    - Both years exceed this limit by a wide margin, especially 2022, which is over **11 times** the safe level.
3. **Possible Reasons for the Increase:**
    - **Urban growth** and infrastructure expansion.
    - **Increased vehicle count** and private car dependency.
    - **Decreased green cover** or ineffective urban planning.
    - Post-COVID **economic reopening** might have led to rapid industrial emissions.

Environmental Concern:

This trend suggests a **worsening air quality** situation in Kolkata, contradicting global efforts toward air pollution control. It points to the **urgent need for policy-level intervention**, including:

- Promotion of **electric vehicles**.
- **Strict monitoring** of construction dust and industrial waste.
- **Enhancing green zones** and public transport access.

Conclusion:

This project helped identify both **temporal (monthly, weekly, yearly)** and **spatial (location-based)** trends in PM2.5 pollution across Kolkata. It shows the importance of **seasonal preparedness**, **traffic regulation**, and **localized green policies**.

Future Scopes:

- Forecasting future PM2.5 using time-series models (ARIMA, LSTM).
- Correlating with weather (temperature, humidity).
- Including health impact metrics (hospitalization rates).