Air Quality Analysis of Bhubaneswar (2024)

A Seasonal Study of Pollution Trends and Public Health Impact

```
By: Abhinav Anand
Date: August 2025
import pandas as pd
import matplotlib.pyplot as plt
# Load the dataset
df = pd.read_csv("Raw_data_1Day_2024_site_5943_Patia_Bhubaneswar_OSPCB_1Day.csv")
# Backup original
df raw = df.copy()
# Drop fully empty columns
df.dropna(axis=1, how='all', inplace=True)
# Fill missing numeric values with column mean
df.fillna(df.select_dtypes(include='number').mean(), inplace=True)
# Parse datetime
df['Timestamp'] = pd.to_datetime(df['Timestamp'])
# Extract Month
df['Month'] = df['Timestamp'].dt.month
```

Abstract

This project analyzes daily air quality data from Bhubaneswar (2024) to uncover pollution trends across PM2.5, PM10, Ozone, and Temperature. It reveals how pollution levels spike during winter months—posing health risks for the public. The analysis involves data cleaning, monthly aggregation, safe limit comparison, and visualization.

Driven by a personal desire to help those around me live healthier, happier lives, this project combines technical insight with emotional purpose—transforming data into awareness.

Technologies Used

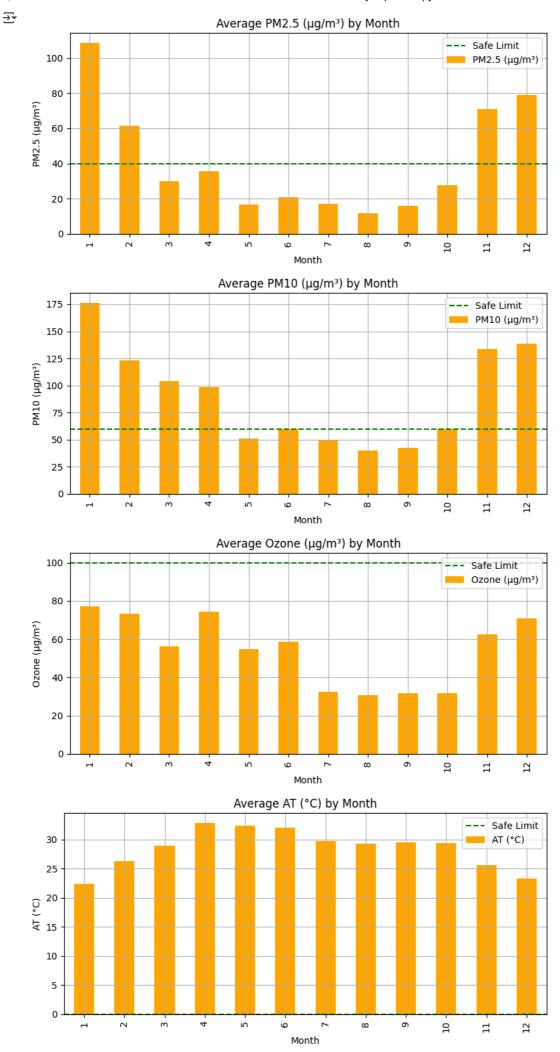
- Python (Pandas, Matplotlib)
- Google Colab Notebook
- · Data Cleaning, Aggregation, Visualization
- Real-world air quality dataset (OSPCB, 2024)

```
# List of pollutants
pollutants = ['PM2.5 (μg/m³)', 'PM10 (μg/m³)', 'Ozone (μg/m³)', 'AT (°C)']
safe_limits = {'PM2.5 (μg/m³)': 40, 'PM10 (μg/m³)': 60, 'Ozone (μg/m³)': 100}

for pollutant in pollutants:
    monthly_avg = df.groupby('Month')[pollutant].mean()

    plt.figure(figsize=(8, 4))
    monthly_avg.plot(kind='bar', color='orange')
    plt.axhline(y=safe_limits.get(pollutant, 0), color='green', linestyle='--', label='Safe Limit')
    plt.title(f'Average {pollutant} by Month')
    plt.ylabel('Month')
    plt.ylabel(pollutant)
    plt.legend()
    plt.grid(True)
    plt.tight_layout()
    plt.show()
```





```
# Count days exceeding safe limits
for pollutant, limit in safe_limits.items():
    count = (df[pollutant] > limit).sum()
    print(f"{pollutant} exceeded safe limit on {count} days.")

PM2.5 (μg/m³) exceeded safe limit on 143 days.
    PM10 (μg/m³) exceeded safe limit on 226 days.
    Ozone (μg/m³) exceeded safe limit on 11 days.
```

Summary of Findings

- January, November, and December show dangerously high levels of PM2.5 and PM10.
- PM2.5 crossed safe levels on 143 days, and PM10 on 226 days in 2024—alarming.
- Ozone levels are mostly within safe range, exceeding only 11 times.
- There is a clear seasonal pattern, with colder months having significantly worse air quality.
- Temperature and pollution levels are inversely related, suggesting winter air stagnation may trap pollutants.

o Implications

- People with asthma, children, and elders are at higher risk in winter.
- · Awareness campaigns and masks can help.
- Authorities can plan for targeted action (air purifiers in schools, traffic restrictions in Jan/Nov/Dec).

Reflection

This project taught me something that I didn't fully realize before:

We breathe numbers—but those numbers carry our health, our comfort, and even our joy.

I started this project wanting to make life better for the people around me.

Now I see what they're breathing every day.

And because I saw it, I can do something about it.

Even if it's small—like sharing this, raising awareness, or inspiring someone to act.

This is how we bring happiness, one insight at a time.