Team Members: Wenyu Lu, Jiaqi Zhu, Minseo Kim

Project Title: Air Pollution and Asthma Emergency Department Visits in New York City: A Time-Series and Neighborhood-Level Analysis

---

1. Motivation & Background

Air pollution is a major public health concern in urban environments, and New York City (NYC) consistently reports high rates of respiratory illnesses, particularly asthma. Fine particulate matter (PM2.5) is known to trigger asthma symptoms and increase emergency department (ED) visits. Understanding how air quality fluctuations relate to asthma outcomes can inform public health interventions and guide city-level environmental policies.

This topic is not only meaningful for public health but also closely connected to our lived environment. The data are publicly available, well-structured, and suitable for the types of modeling and visualization.

This project will explore how PM2.5 levels vary across time and NYC boroughs, and whether increases in pollution are associated with higher asthma ED visit rates.

---

2. Data Sources

We will use two publicly available datasets from NYC government sources:

(A) Air Quality (PM2.5) – NYC Open Data

- Contains daily or monthly PM2.5 concentrations
- Available by citywide average and sometimes by borough
- Well structured and easy to clean
- URL: NYC Open Data → “Air Quality: PM2.5” dataset

(B) Asthma Emergency Department Visits – NYC DOHMH (Health Department)

- Contains borough-level monthly asthma ED visit rates
- Includes age-standardized visit rates

- Data are consistent, tidy, and human-readable
- URL: NYC Environment & Health Data Portal → “Asthma Emergency Department Visits”

We will merge datasets by month-year and borough.

---

3. Research Questions

Our project will address the following questions:

1. How have PM2.5 levels and asthma ED visit rates changed over time in NYC (2015–2023)?
2. Do boroughs differ in both PM2.5 concentration and asthma ED visit rates?
3. Is higher PM2.5 associated with higher asthma ED visit rates after adjusting for seasonality and borough?
4. Does season modify the association between PM2.5 and asthma ED visits?

These questions allow for both exploratory data analysis and regression modeling.

---

4. Planned Data Cleaning & Structure

We will produce one tidy merged dataset:

- Convert dates to year-month format
- Merge PM2.5 and asthma datasets by borough and month
- Create categorical variables:
    - borough (factor)
    - season (Winter, Spring, Summer, Fall)
- Optional: merge temperature data

- Check for missing values and address them appropriately

The final dataset will contain:

month, borough, pm25, asthma_ed_rate, season, temperature

---

5. Planned Analyses

A. Exploratory Data Analysis (EDA)

- Time-series plots of PM2.5 by borough（Jiaqi）
- Time-series plots of asthma ED visit rates（Jiaqi）
- Scatterplots comparing PM2.5 vs asthma ED rates (Bruce)
- Seasonal trends (Bruce)
- Borough comparison using faceting (Brenda)
- Heatmaps (month × borough) (Brenda)

B. Statistical Modeling

We will fit at least one multivariable regression model:

Model 1: Linear Regression

Model 2: Poisson/Negative Binomial Regression

Model 3: Interaction Model

We will evaluate model assumptions, present effect estimates, and interpret public health implications.

---

6. Expected Deliverables

- GitHub Repo with R Project, data, scripts, report

- A polished website (GitHub Pages) presenting the analysis
- Clear and well-designed visualizations
- Final report (HTML)
- 2 minute screencast summarizing results