University of Milano-Bicocca
Master Degree in Data Science
Academic Year 2023/2024

Air Pollution and Public Health in New York City

Presented by:
Paola Maria Cavana

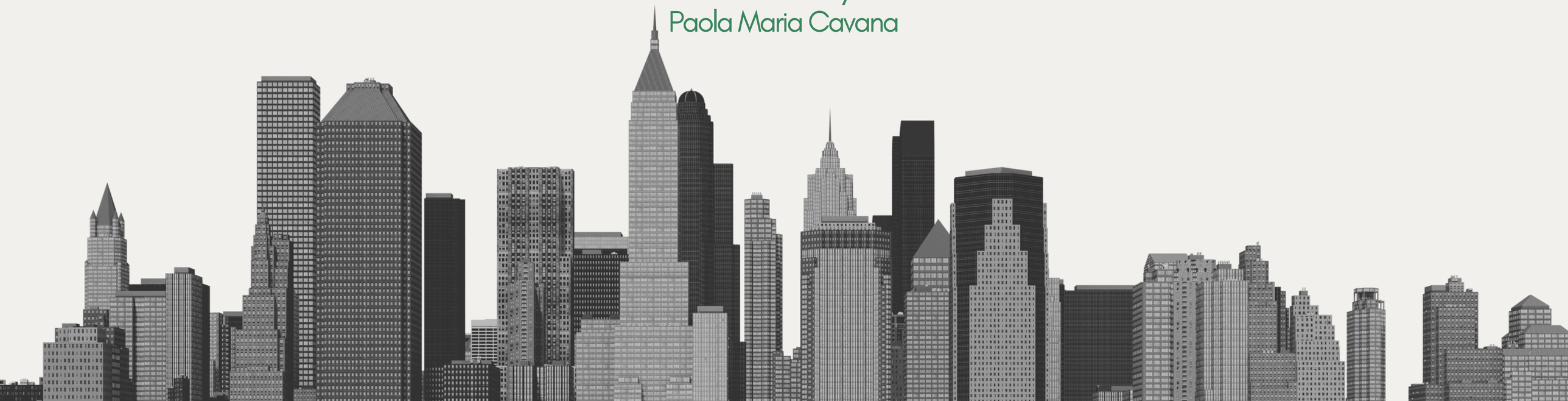


Table of contents

What is air pollution?

Air pollution is the presence of harmful or excessive quantities of substances in the air, including gases, particulates, and biological molecules. It has significant impacts on **human health**, **animals**, and the **environment**. For humans, air pollution causes **respiratory diseases** like asthma and bronchitis, contributes to **cardiovascular problems**, and is linked to premature death and reduced life expectancy. The main sources of pollution include **industrial activities**, **vehicular emissions**, **agricultural activities** and **household heating**.

Ethical and social implications

Air pollution disproportionately affects various communities, often impacting low-income and minority groups, children, the elderly, and individuals with preexisting health conditions more severely.

These communities are often situated in areas with elevated pollution levels as a result of their proximity to:

Industrial zones

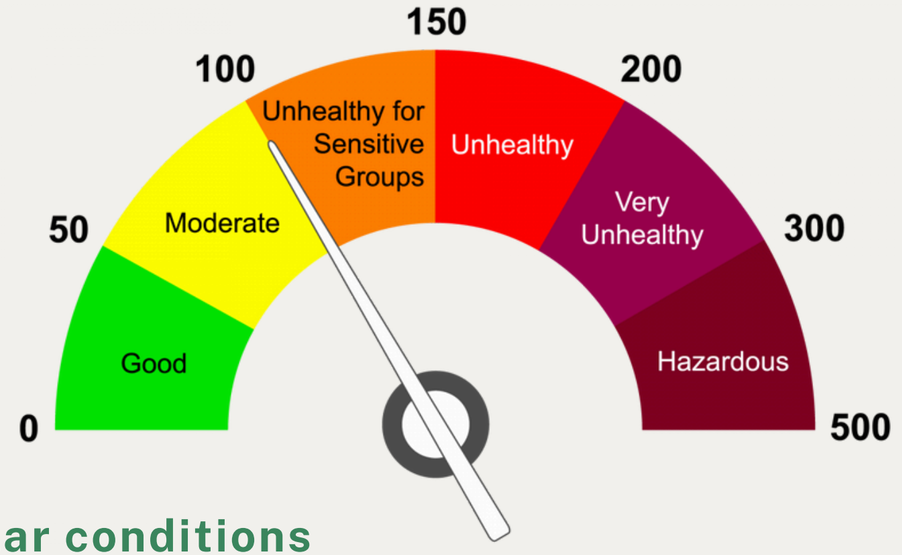
Highways

Waste disposal sites

Selection of indicators

Air quality indicators:

- Concentrations of pollutants such as Black Carbon, PM2.5, NO2, O3, and NO
- Air Quality Index (AQI)



Health indicators:

- Hospitalization rates for respiratory and cardiovascular conditions
- Mortality rates
- Emergency department visits related to asthma

Selected Dataset

The analysis of air pollution and its health impacts in New York City relies on several key datasets.

Air Quality Data:

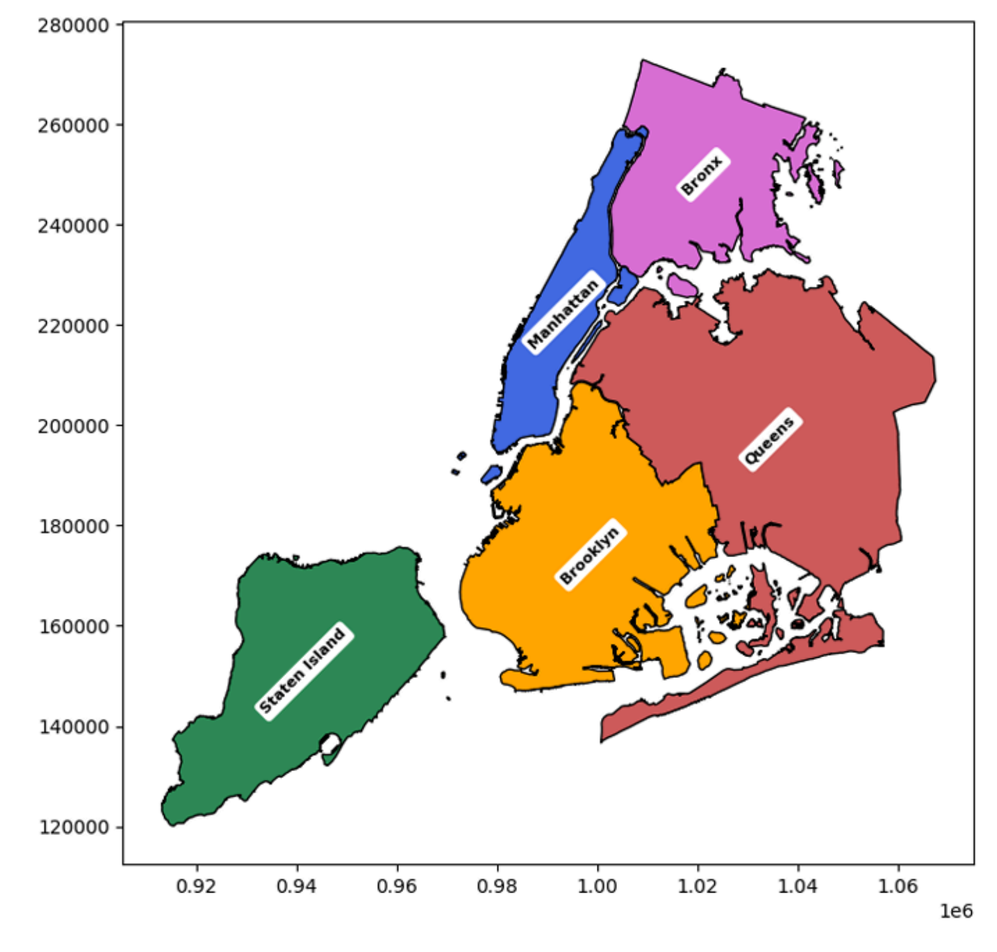
- Source: NYC.GOV, World Air Quality Historical Database and NYC OpenData.
- Pollutants monitored: Black Carbon, PM2.5, NO2, O3, NO.
- Data collection methods: Fixed monitoring stations, mobile sensors, and satellite observations.

Health Data:

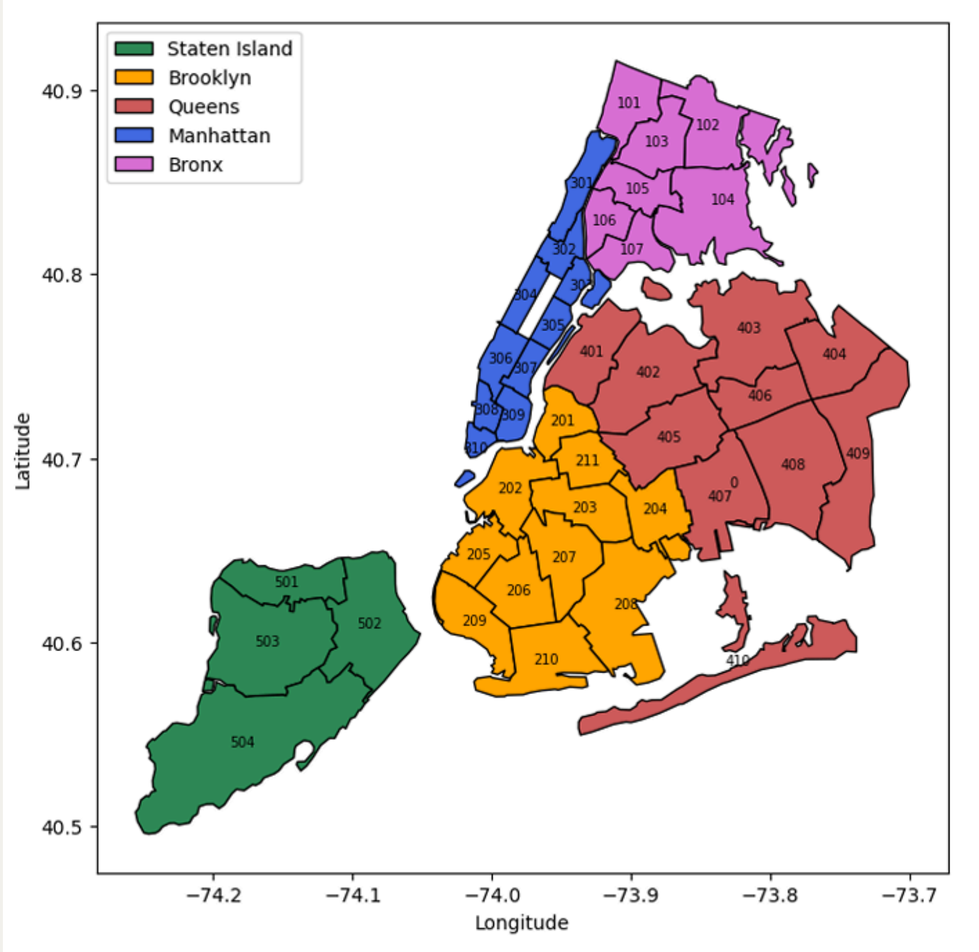
- Source: Environmental & Health Data Portal.
- Indicators tracked: Hospitalization rates, mortality rates, and emergency department visits for asthma and other conditions.
- Data collection methods: Medical records, health surveys, and public health databases.

Spatial Analysis

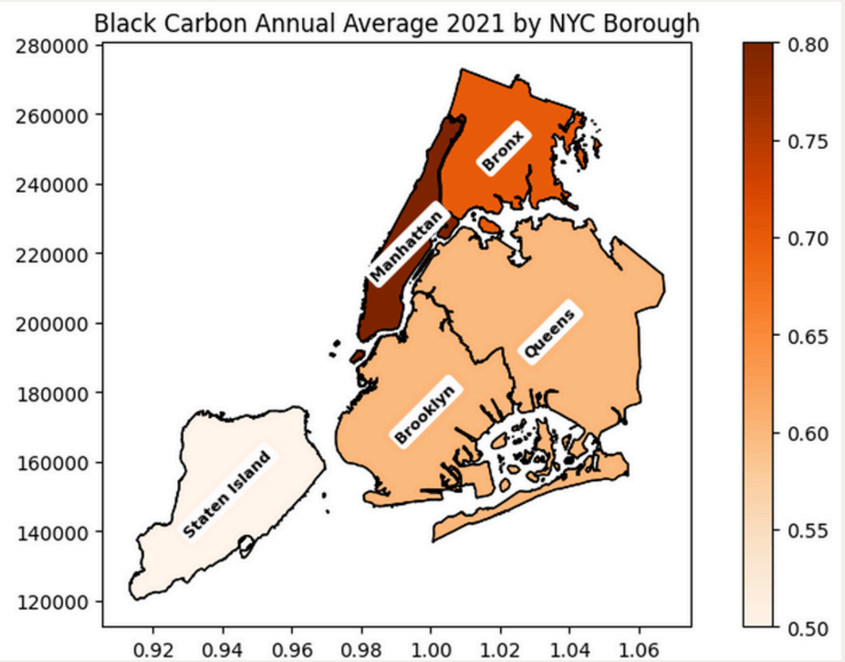
Two distinct geographical mappings:

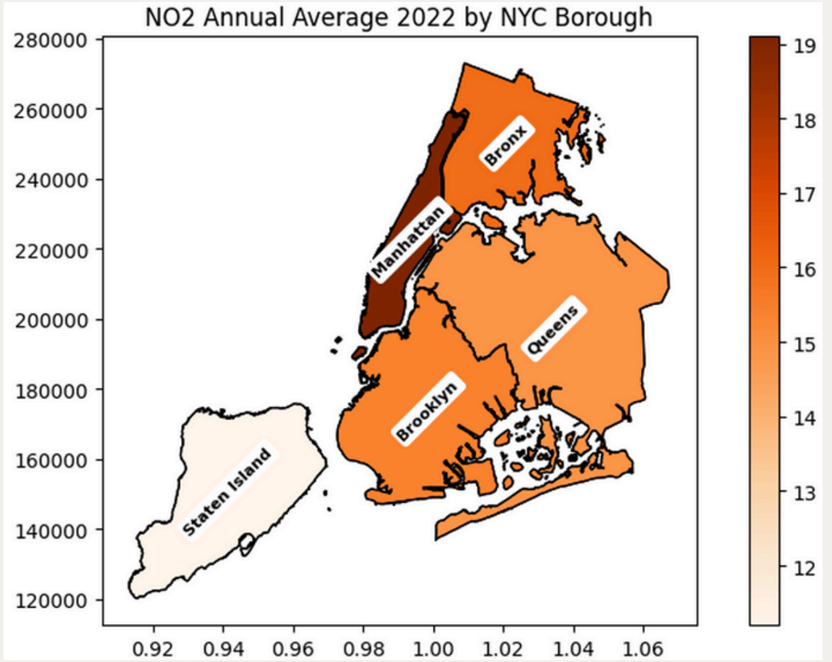


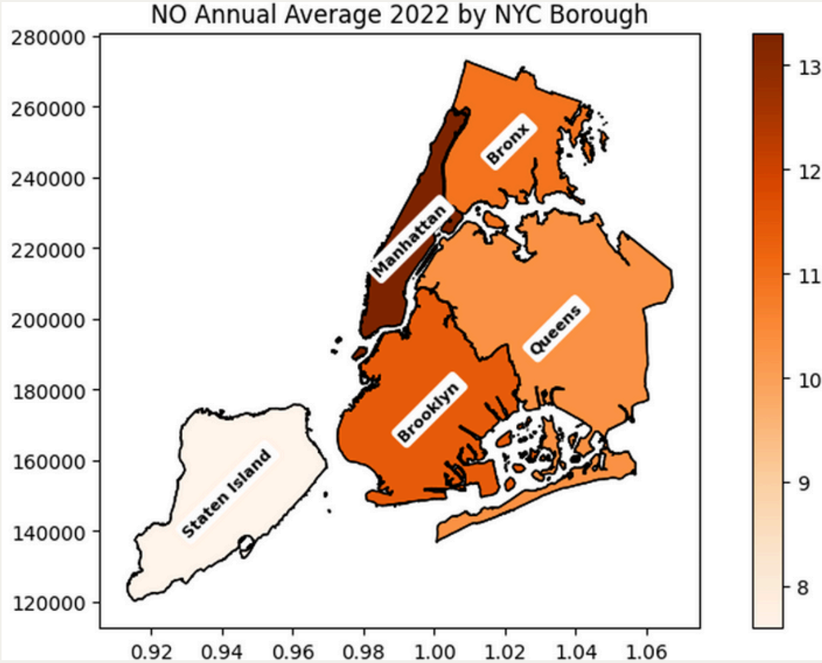
NYC Borough map

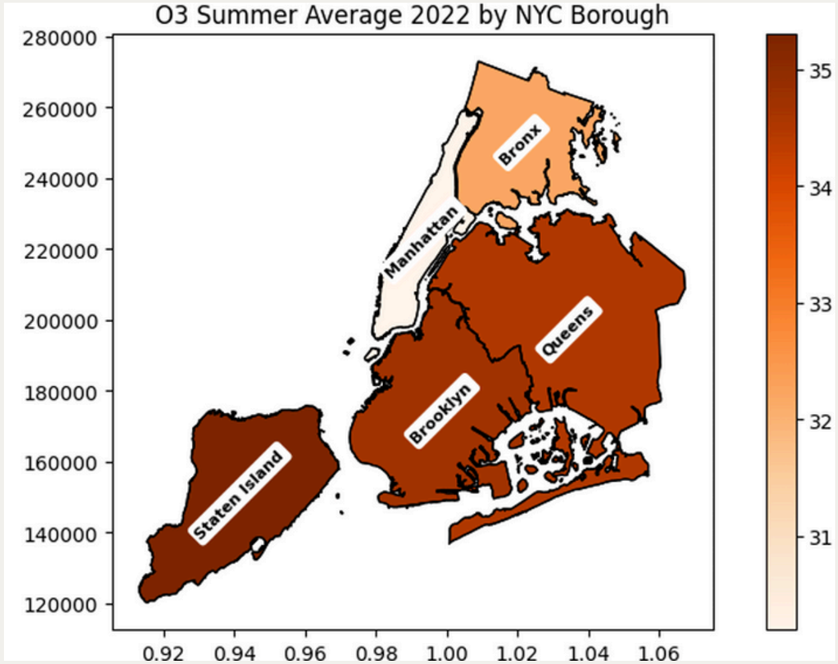


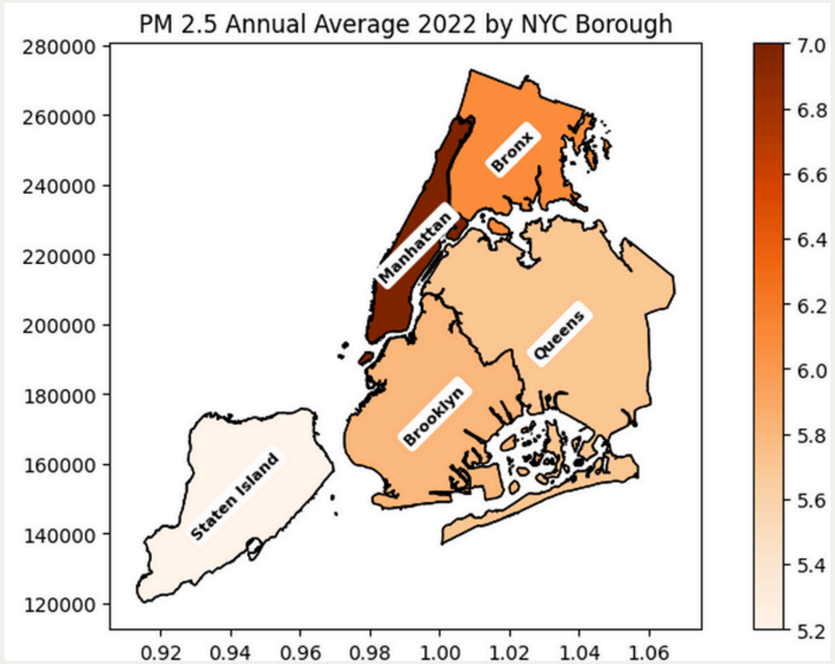
NYC UHF42 map

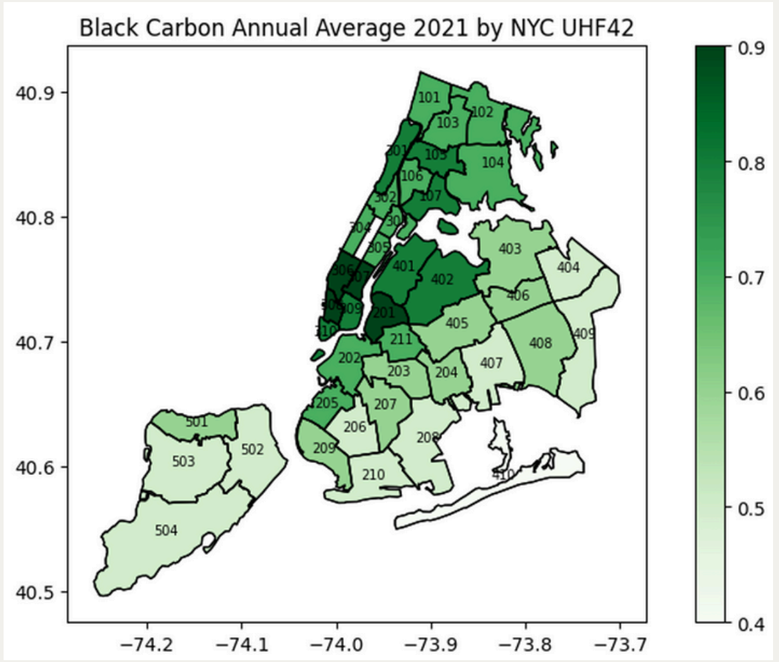

Black Carbon Annual Average 2021 by NYC Borough
Bronx
Manhattan
Queens
Brooklyn
Staten Island
280000
260000
240000
220000
200000
180000
160000
140000
120000
0.92
0.94
0.96
0.98
1.00
1.02
1.04
1.06
0.80
0.75
0.70
0.65
0.60
0.55
0.50

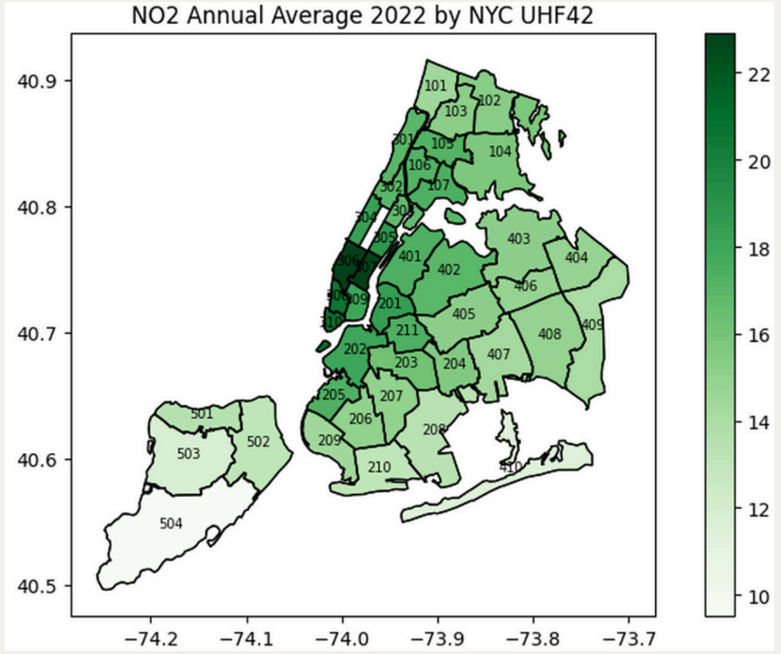

NO2 Annual Average 2022 by NYC Borough
Bronx
Manhattan
Queens
Brooklyn
Staten Island
280000
260000
240000
220000
200000
180000
160000
140000
120000
0.92
0.94
0.96
0.98
1.00
1.02
1.04
1.06
19
18
17
16
15
14
13
12

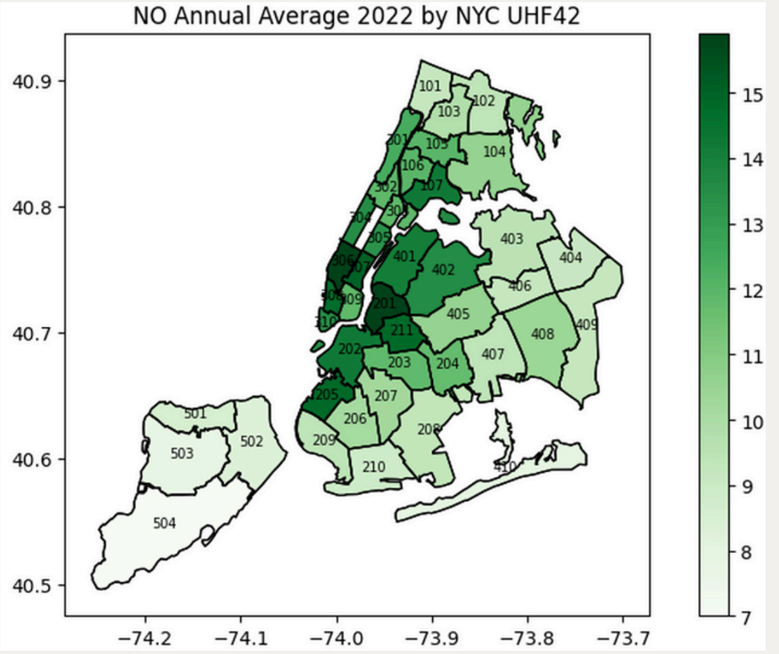

NO Annual Average 2022 by NYC Borough
Bronx
Manhattan
Queens
Brooklyn
Staten Island
280000
260000
240000
220000
200000
180000
160000
140000
120000
0.92
0.94
0.96
0.98
1.00
1.02
1.04
1.06
13
12
11
10
9
8

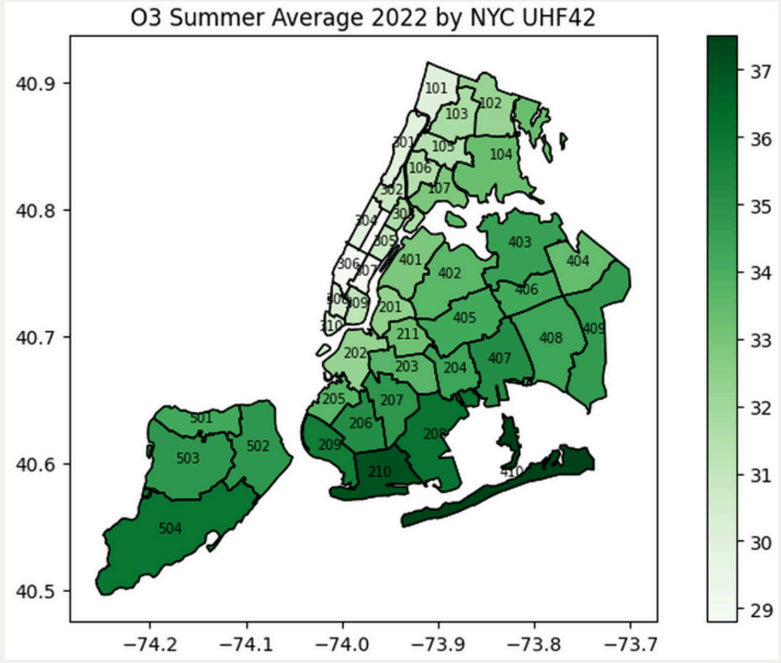

O3 Summer Average 2022 by NYC Borough
Bronx
Manhattan
Queens
Brooklyn
Staten Island
280000
260000
240000
220000
200000
180000
160000
140000
120000
0.92
0.94
0.96
0.98
1.00
1.02
1.04
1.06
35
34
33
32
31

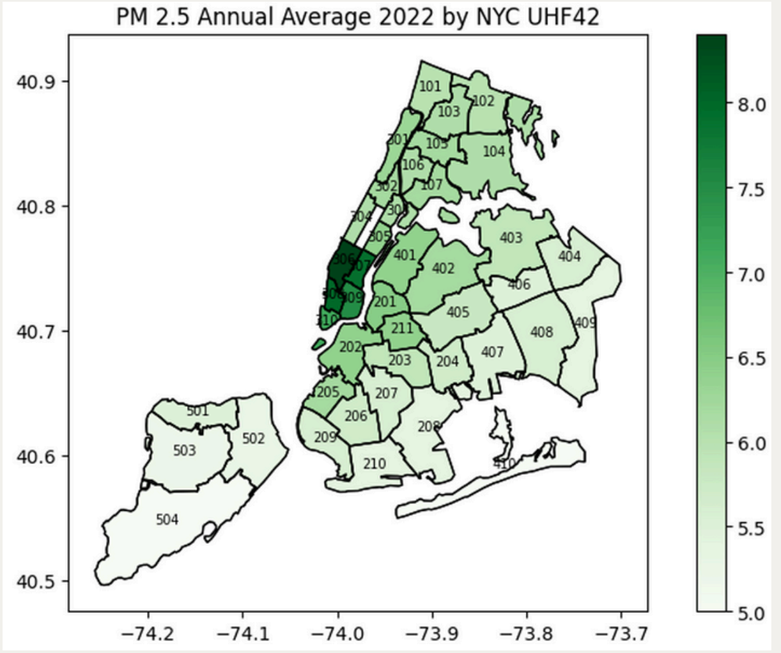

PM 2.5 Annual Average 2022 by NYC Borough
Bronx
Manhattan
Queens
Brooklyn
Staten Island
280000
260000
240000
220000
200000
180000
160000
140000
120000
0.92
0.94
0.96
0.98
1.00
1.02
1.04
1.06
7.0
6.8
6.6
6.4
6.2
6.0
5.8
5.6
5.4
5.2

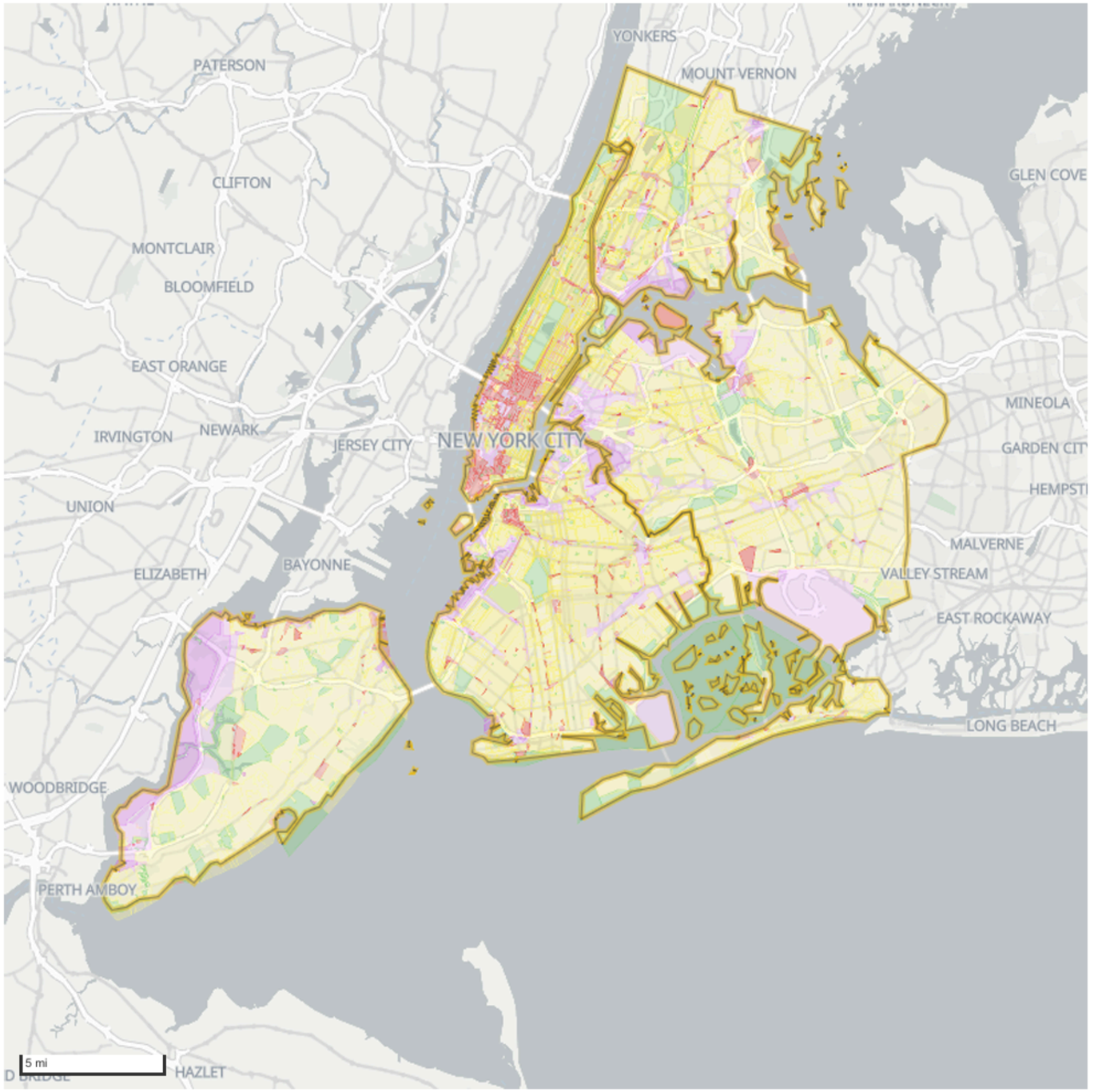

Black Carbon Annual Average 2021 by NYC UHF42


NO2 Annual Average 2022 by NYC UHF42

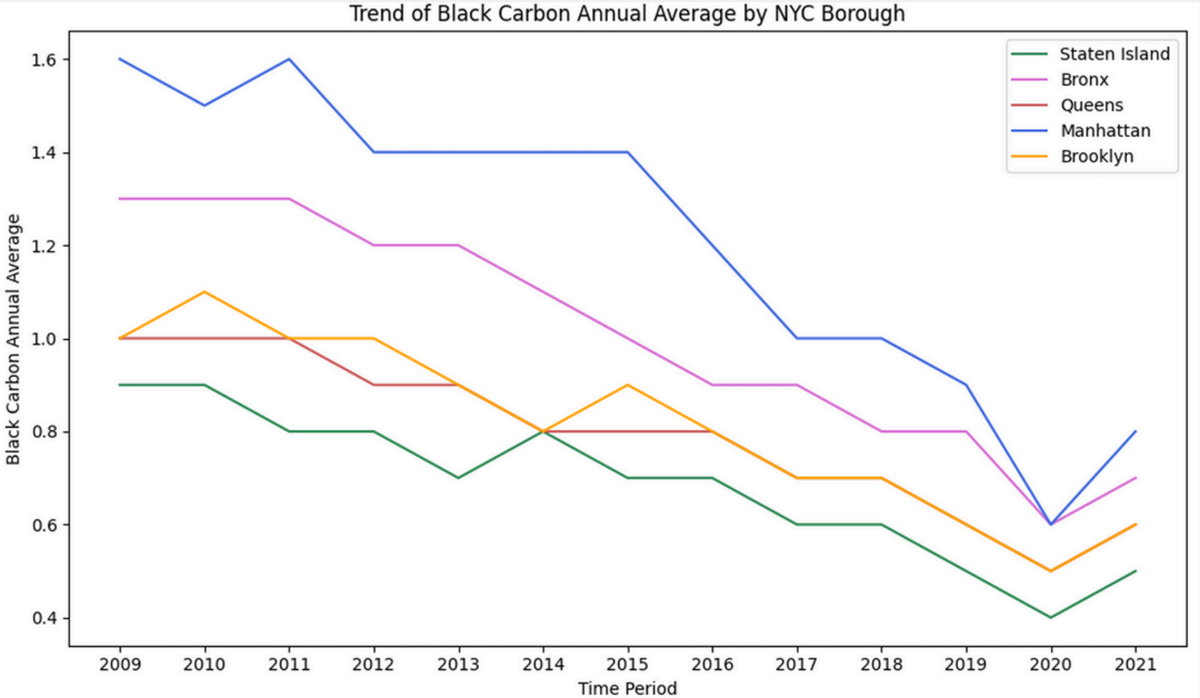

NO Annual Average 2022 by NYC UHF42

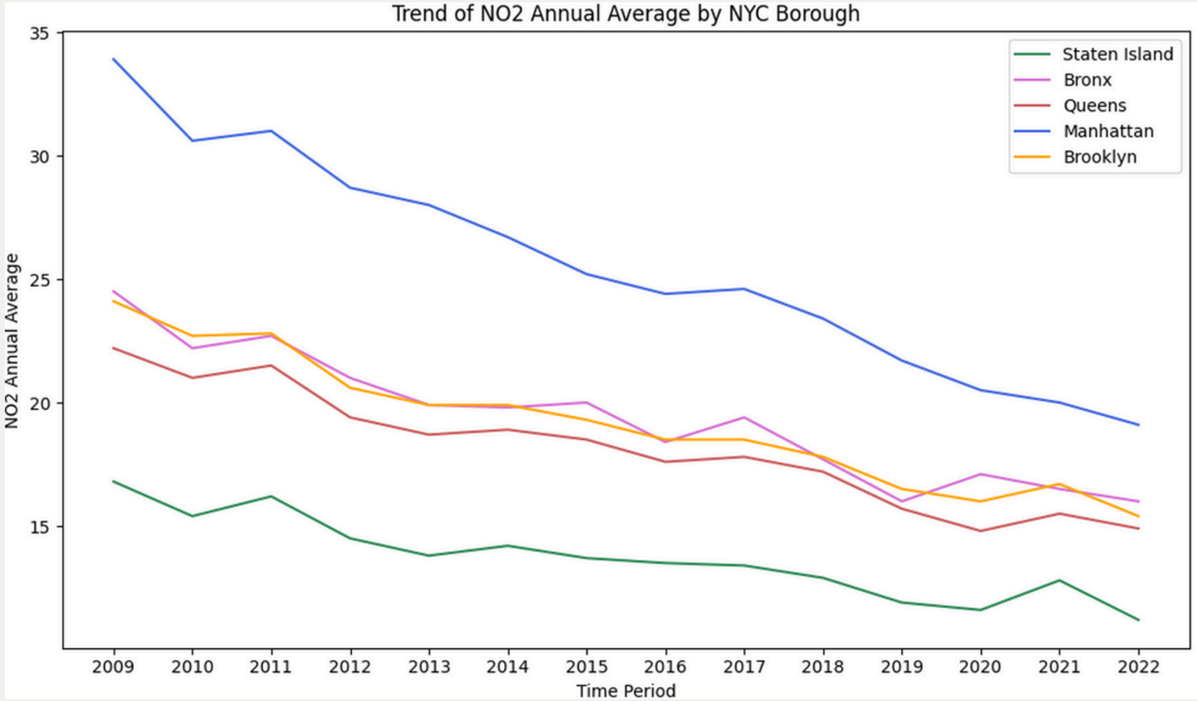

O3 Summer Average 2022 by NYC UHF42


PM 2.5 Annual Average 2022 by NYC UHF42

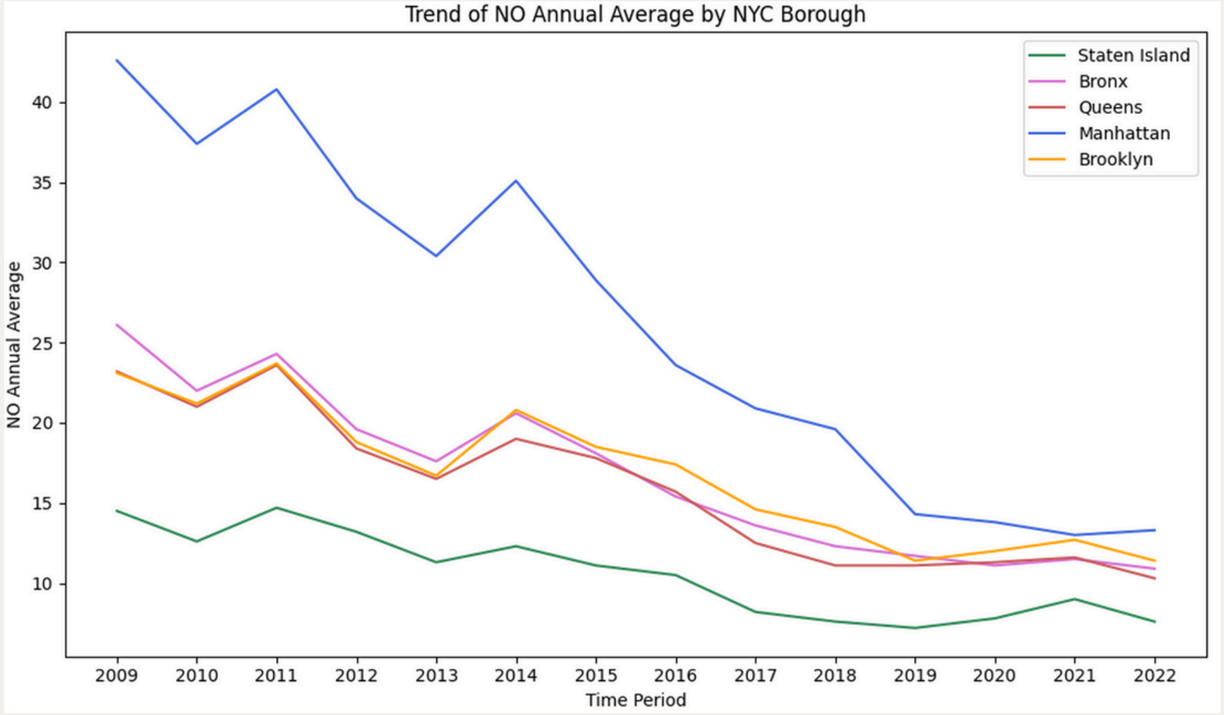

YONKERS
PATERSON
MOUNT VERNON
CLIFTON
GLEN COVE
MONTCLAIR
BLOOMFIELD
EAST ORANGE
MINEOLA
IRVINGTON
NEWARK
JERSEY CITY
NEW YORK CITY
GARDEN CITY
HEMPSTEAD
UNION
MALVERNE
ELIZABETH
BAYONNE
VALLEY STREAM
EAST ROCKAWAY
LONG BEACH
WOODBRIDGE
PERTH AMBOY
HAZLET
5 mi
Zoning and Land Use
Tax Lots
Zoning Districts
Commercial Districts
Manufacturing Districts
Residence Districts
Parks
Battery Park City
Administrative Boundaries
Boroughs

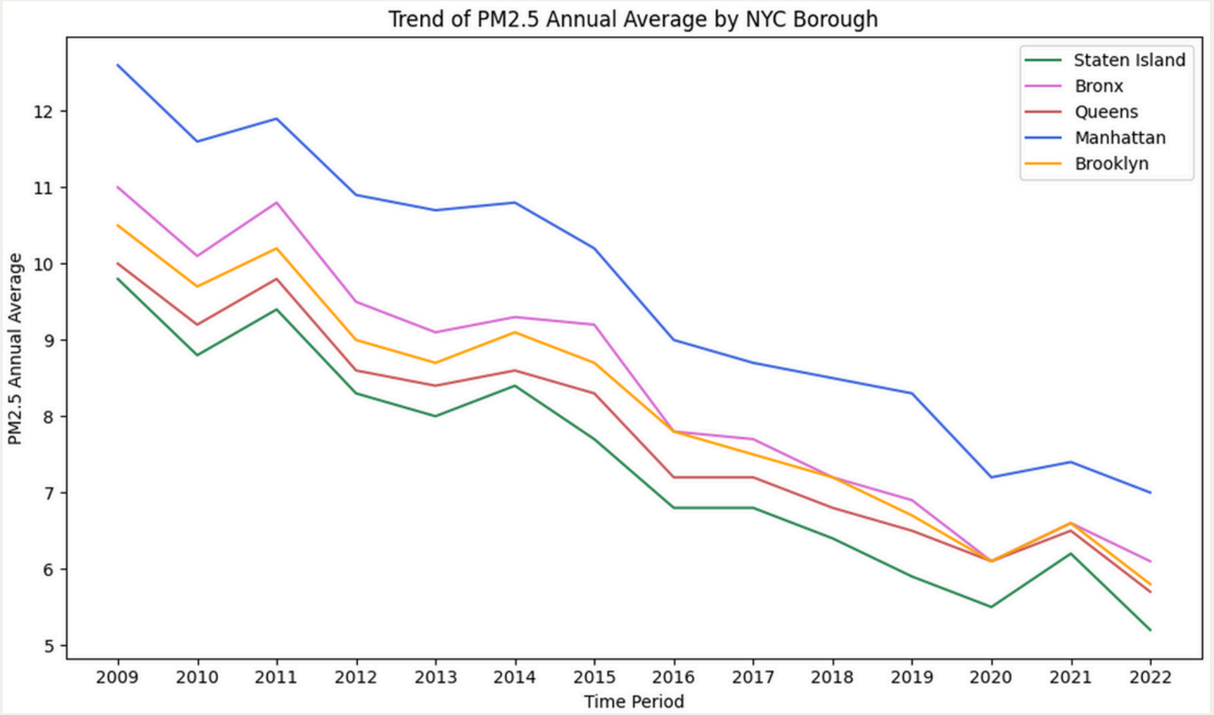
Temporal Analysis



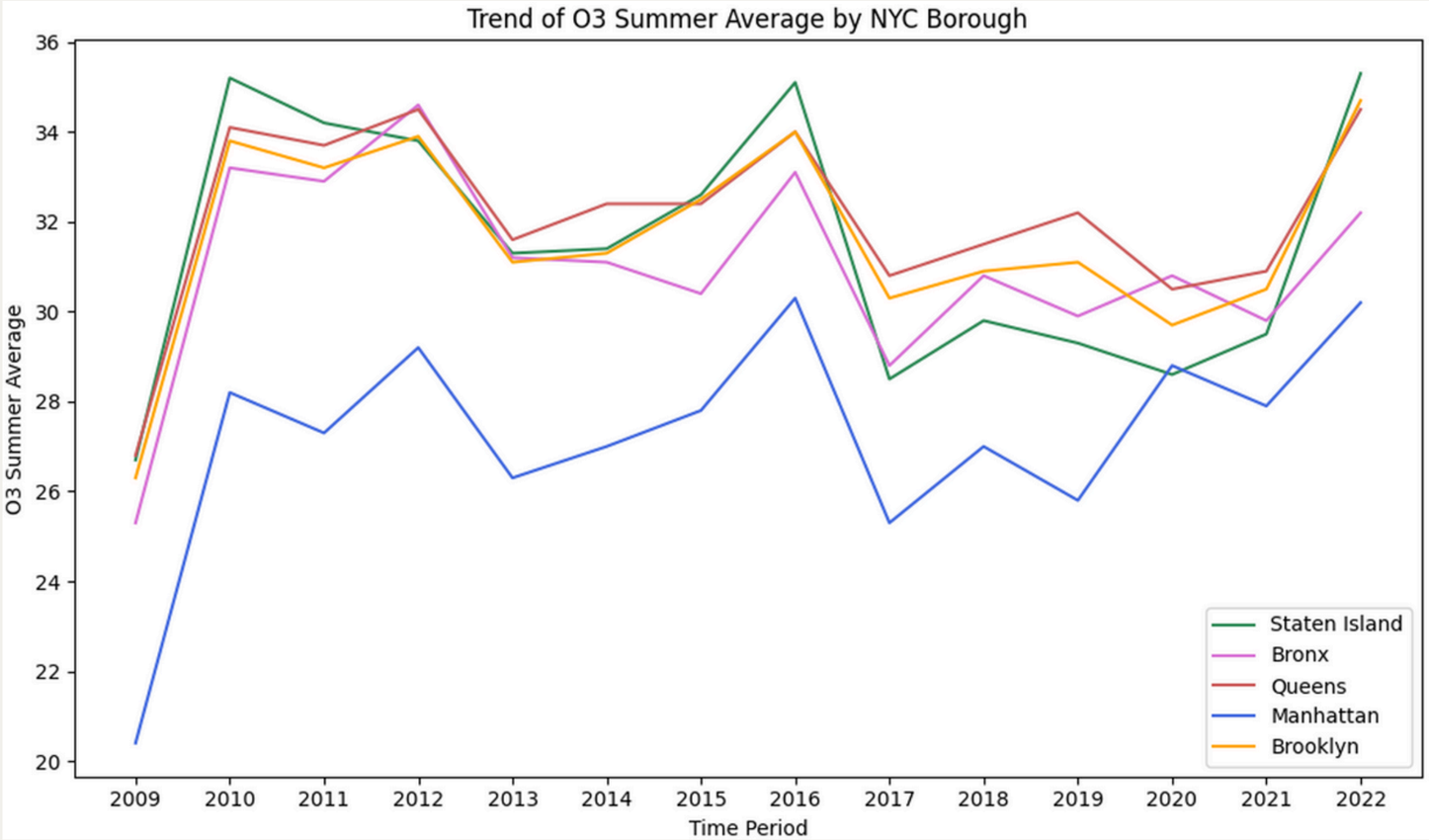


Temporal Analysis

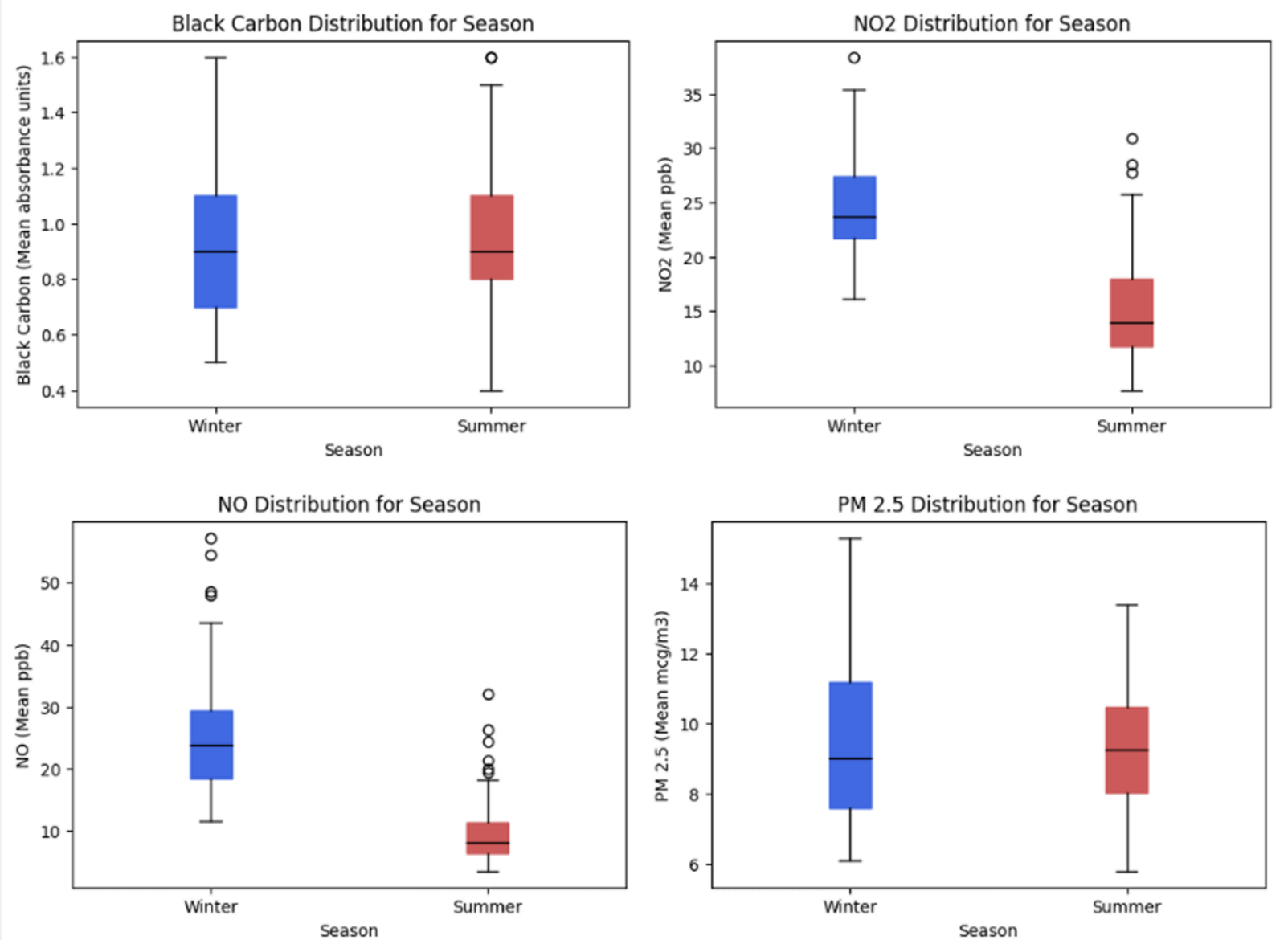




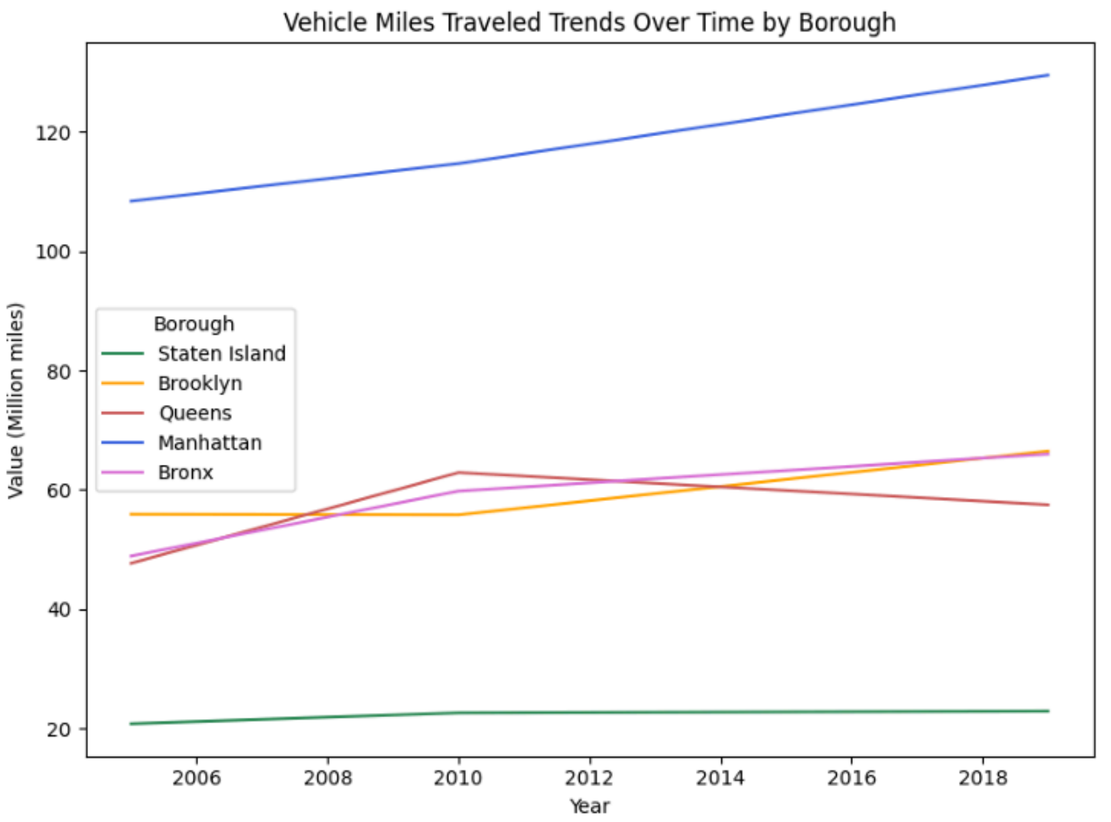
Temporal Analysis



Seasonality Analysis



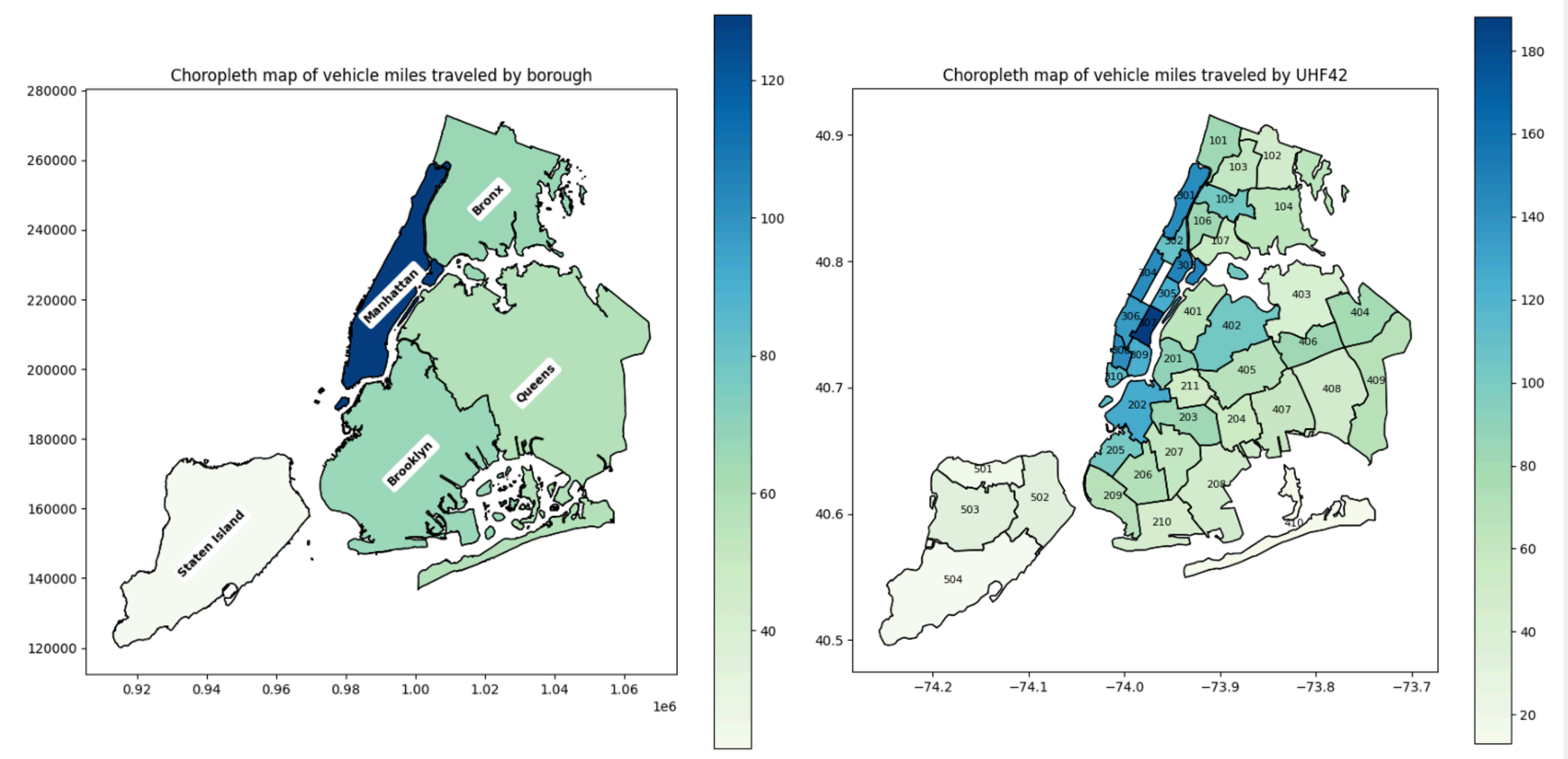
Traffic Analysis



The aggregated data highlights these results:

1. Manhattan: 117.5 million miles per square mile
2. Brooklyn: 59.4 million miles per square mile
3. Bronx: 58.2 million miles per square mile
4. Queens: 56.0 million miles per square mile
5. Staten Island: 22.1 million miles per square mile

Traffic Analysis



Health impact

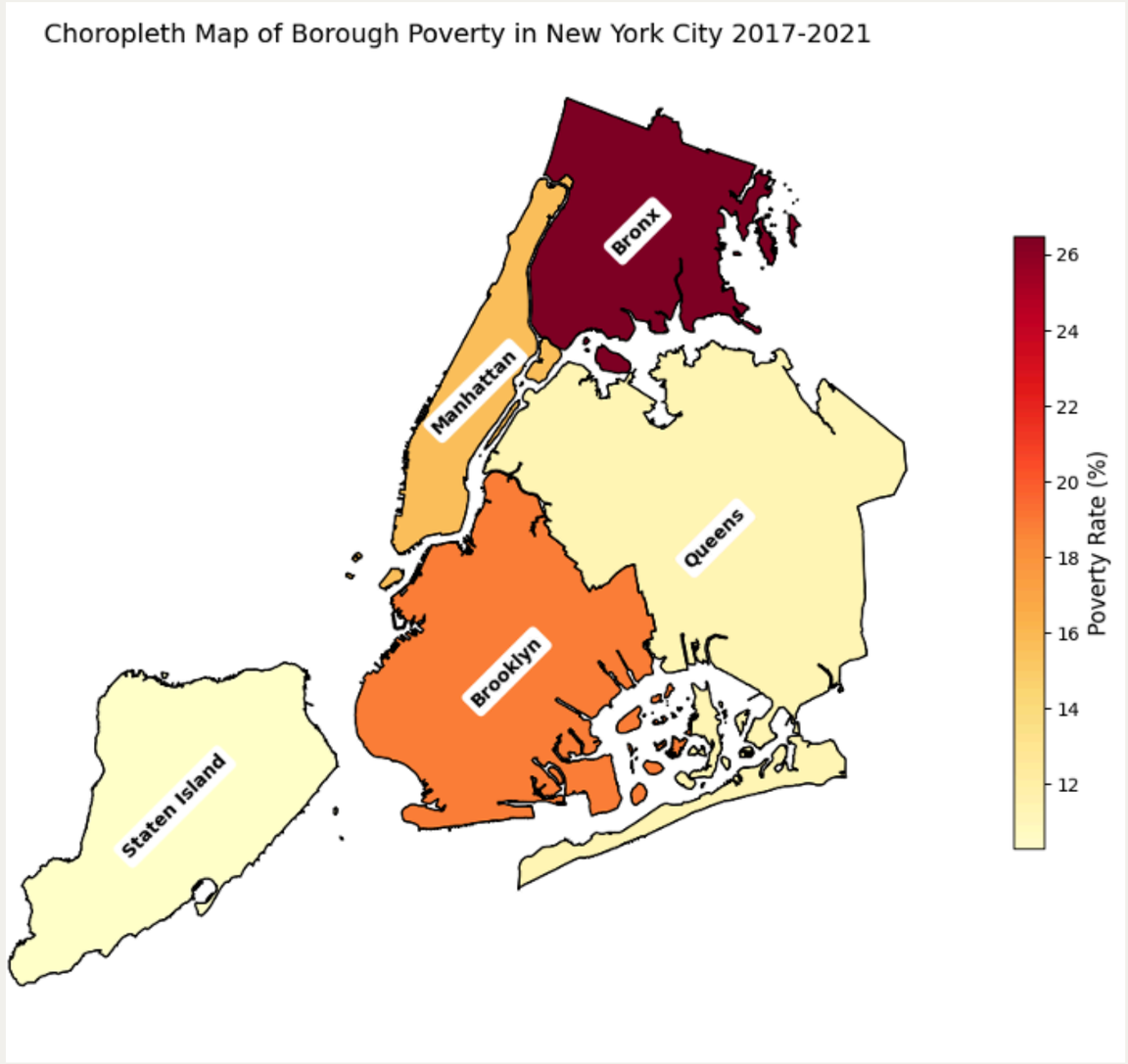
The health impact analysis in this study focuses on the relationship between air pollution and health outcomes across different boroughs of New York City (NYC).
To understand the health impact of air quality and pollution, an examination of the poverty rate of different boroughs and the distribution of health insurance among adults was conducted. These factors are significant because they can influence access to healthcare and the ability to manage pollution-related health conditions.

01. Poverty Rate

02. Health Insurance Coverage

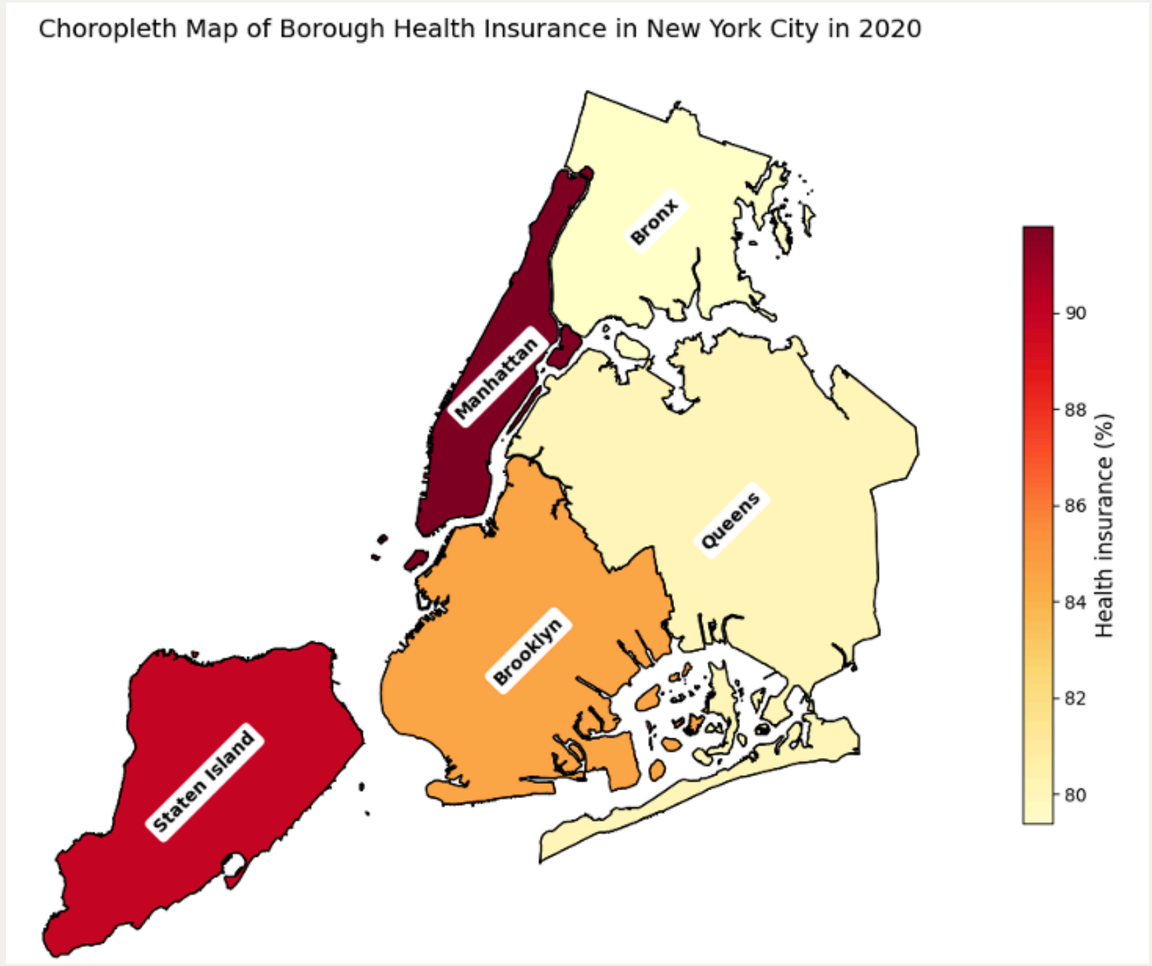
01. Poverty Rate

1. **Bronx: 26.5%**
2. **Brooklyn: 18.8%**
3. **Manhattan: 15.6%**
4. **Queens: 11.4%**
5. **Staten Island: 10.3%**



02. Health Insurance Coverage

1. **Manhattan: 91.8%**
2. **Staten Island: 89.9%**
3. **Brooklyn: 84.6%**
4. **Queens: 80.1%**
5. **Bronx: 79.4%**



Asthma Emergency

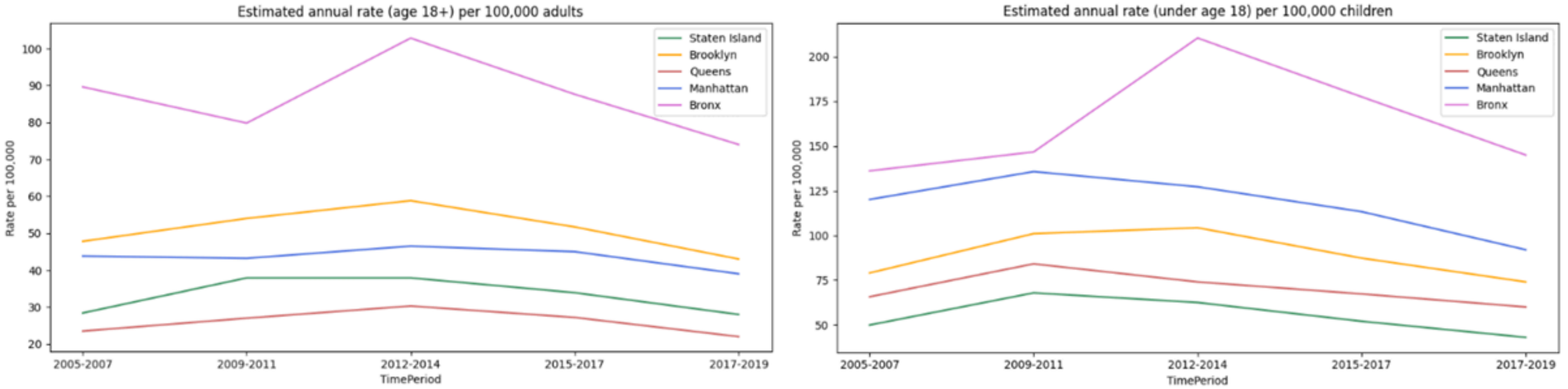


Figure 18: Temporal Trends of Asthma Emergency Rates due to Ozone in NYC Boroughs (2005-2019)

Note: The left plot depicts estimated annual rates per 100,000 adults (age 18+), while the right plot shows rates for children (under age 18).

Asthma Emergency

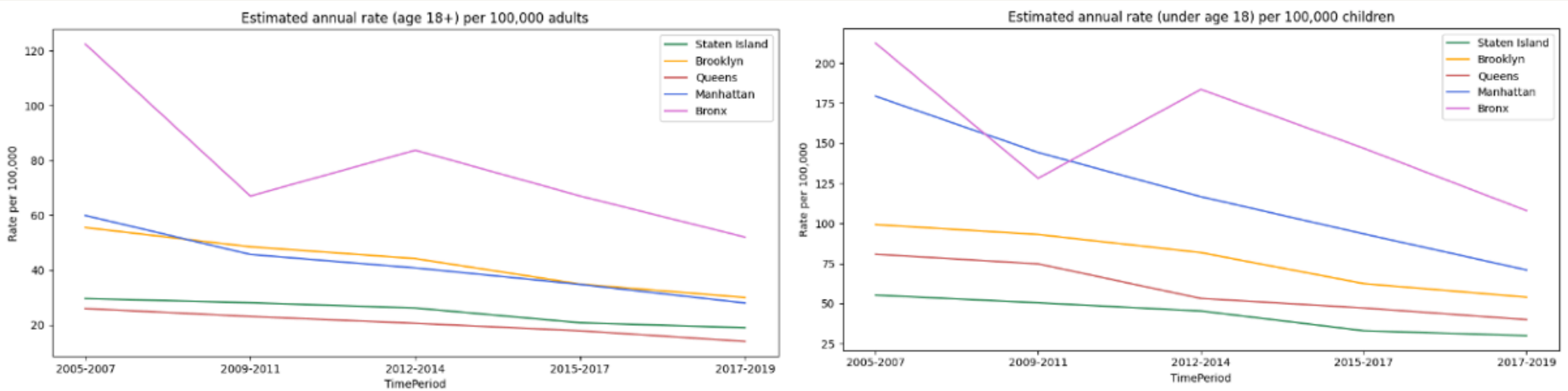


Figure 19: Temporal Trends of Asthma Emergency Rates due to PM2.5 in NYC Boroughs (2005-2019)

Note: The left plot depicts estimated annual rates per 100,000 adults (age 18+), while the right plot shows rates for children (under age 18).

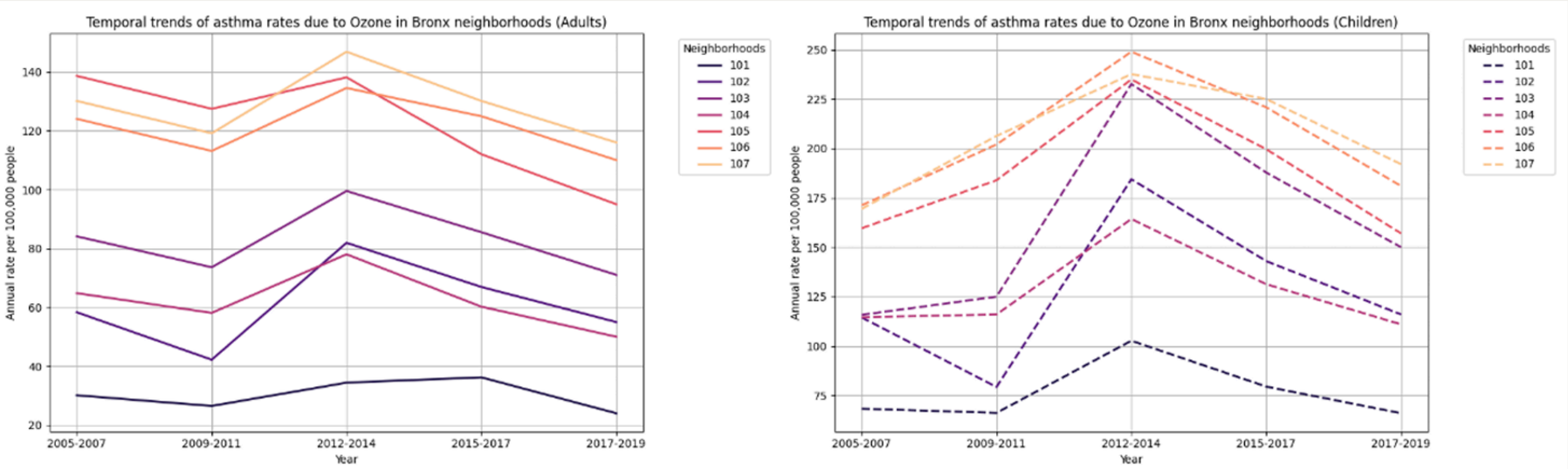


Figure 20: Temporal Trends of Asthma Emergency Rates due to Ozone in Bronx UHF42 (2005-2019)

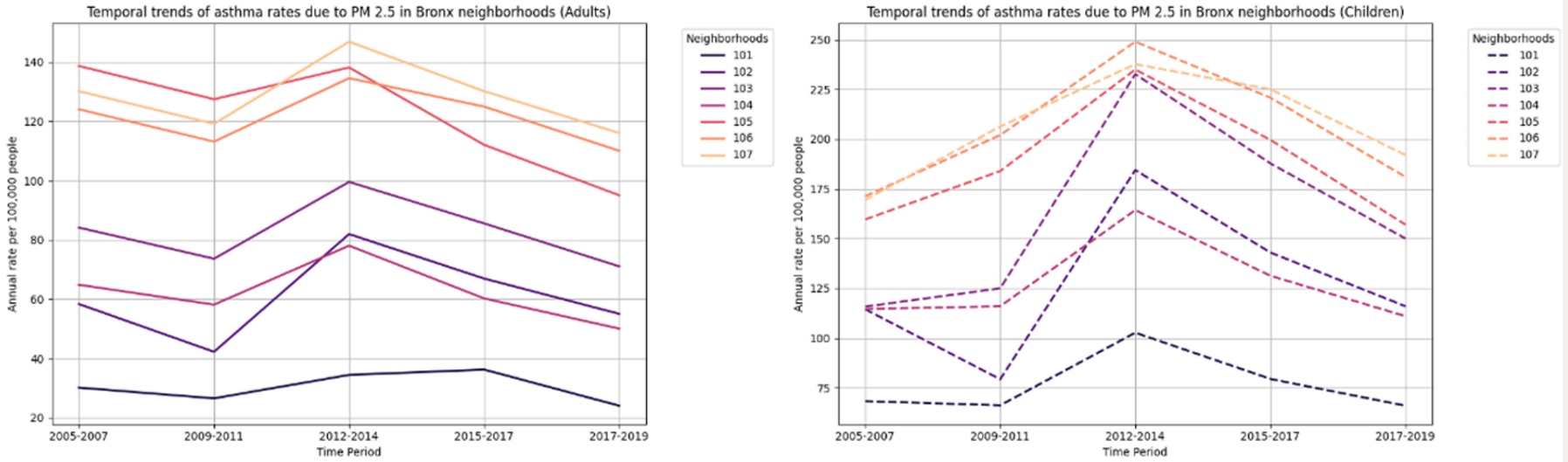
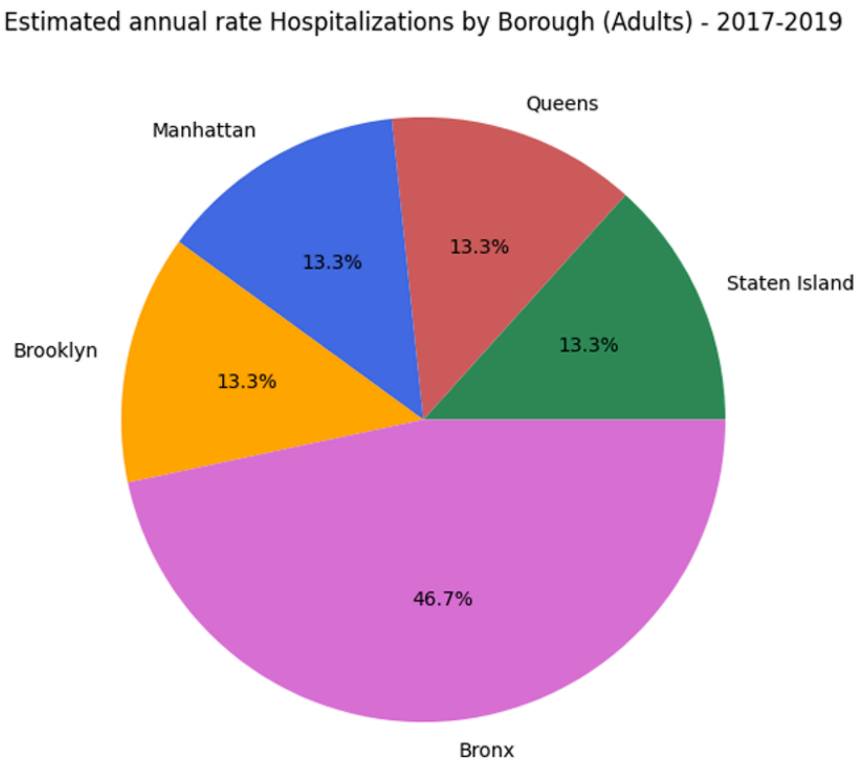


Figure 21: Temporal Trends of Asthma Emergency Rates due to PM2.5 in Bronx UHF42 (2005-2019)

Note: The left plot depicts estimated annual rates per 100,000 adults (age 18+), while the right plot shows rates for children (under age 18).

Hospitalization



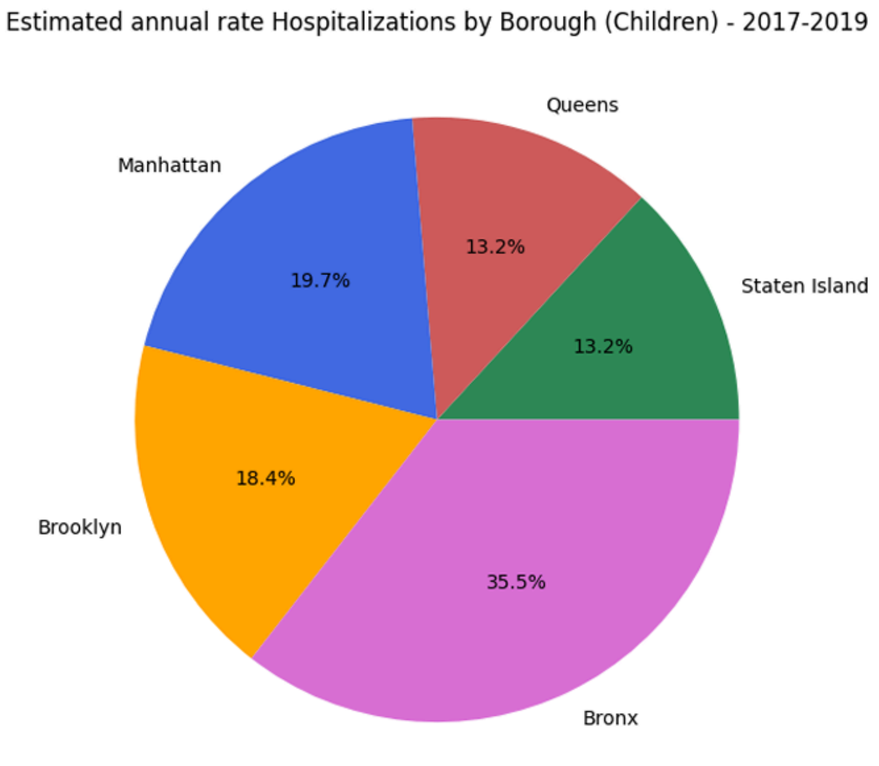


Figure 22: Estimated Asthma Hospitalizations due to O3 by Borough - 2017-2019

Note: The left plot depicts estimated annual rates per 100,000 adults (age 18+), while the right plot shows rates for children (under age 18).

Hospitalization

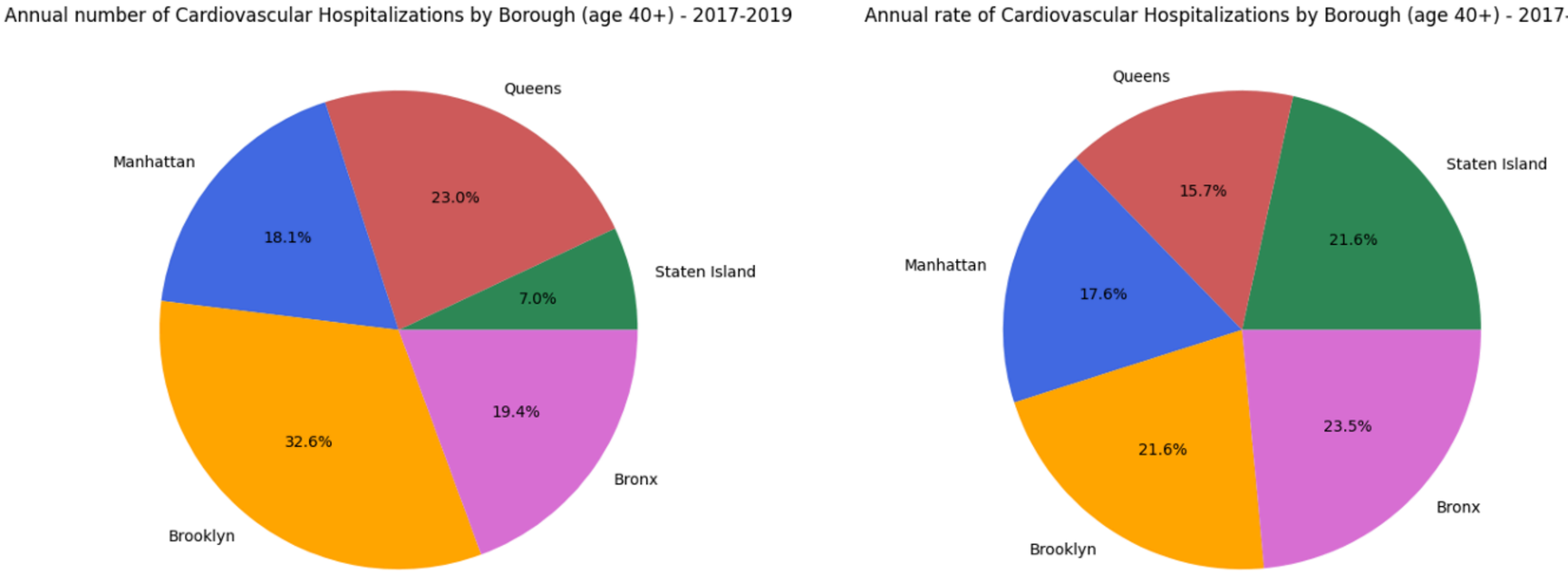
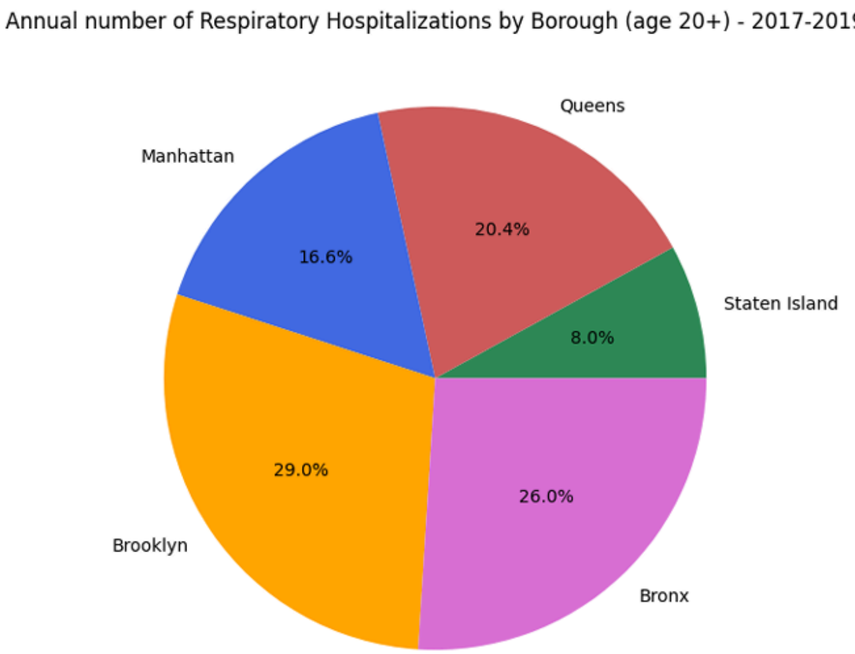


Figure 23: Estimated Cardiovascular Hospitalizations due to PM2.5 by Borough - 2017-2019

Note: The left plot depicts estimated annual number (age 40+), while the right plot shows the estimated annual rates per 100,000 adults (age 40+).

Hospitalization



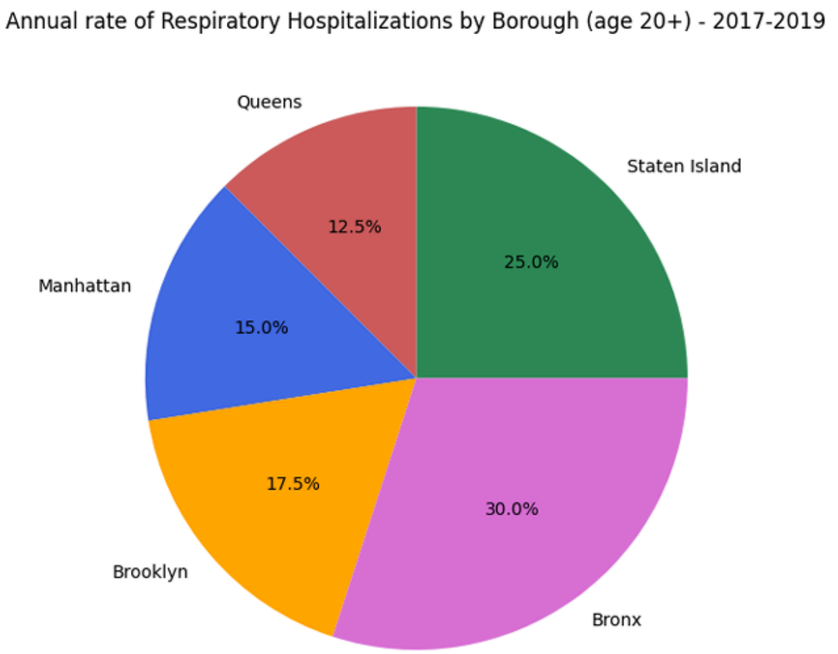
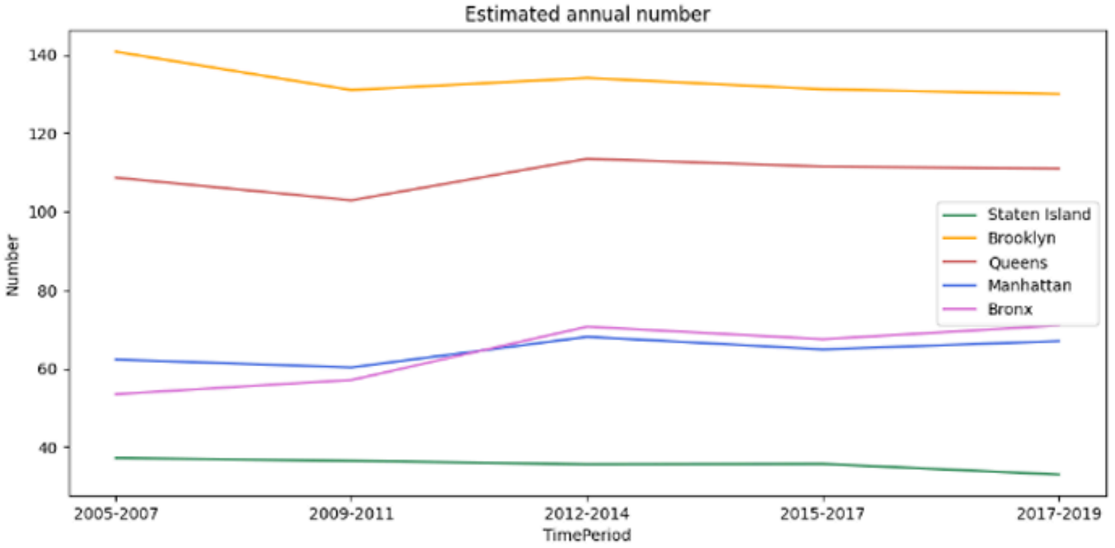


Figure 24: Estimated Cardiovascular Hospitalizations due to PM2.5 by Borough - 2017-2019

Note: The left plot depicts estimated annual number (age 20+), while the right plot shows the estimated annual rates per 100,000 adults (age 20+).

Deaths



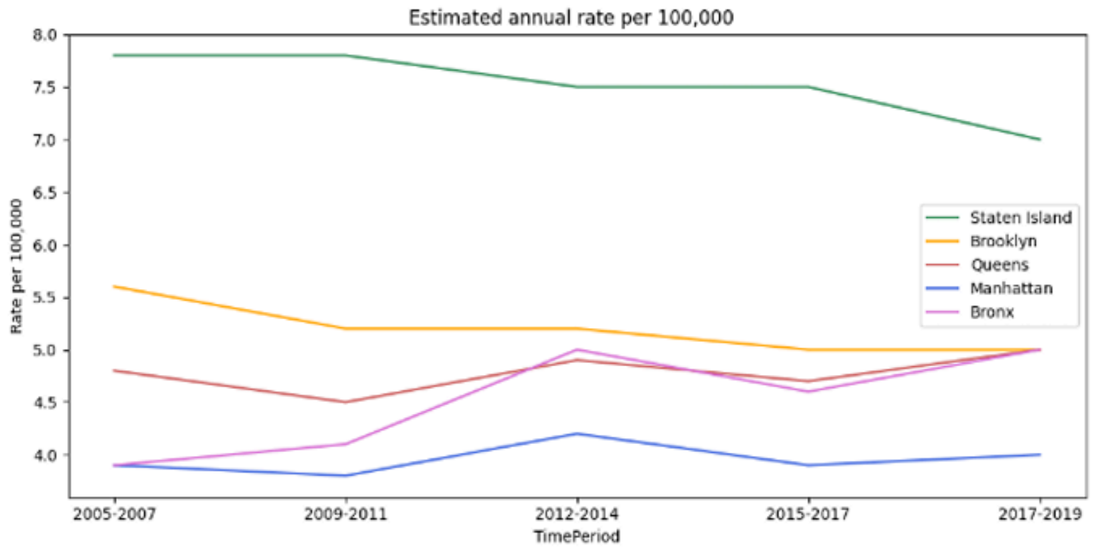
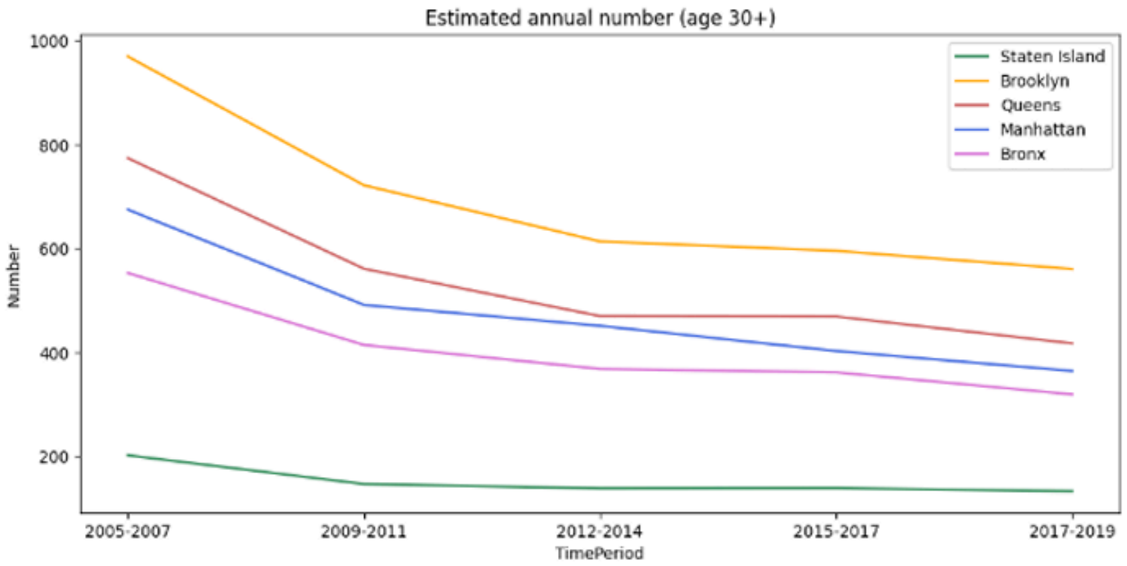


Figure 25: Estimated Deaths due to Ozone by Borough (2005-2019)

Table 1: Deaths due to Ozone by Borough - 2017-2019

TimePeriod	Geography	Estimated annual number	Estimated annual rate
2017-2019	Bronx	71.0	5.0
2017-2019	Brooklyn	130.0	5.0
2017-2019	Manhattan	67.0	4.0
2017-2019	Queens	111.0	5.0
2017-2019	Staten Island	33.0	7.0

Deaths



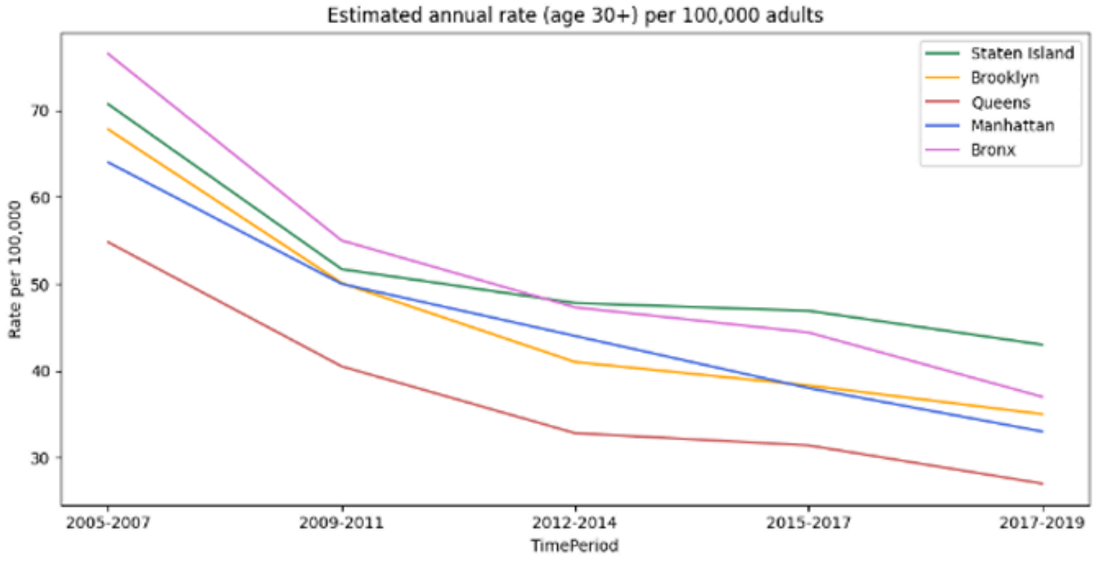


Figure 26: Estimated Deaths due to PM2.5 by Borough (2005-2019)

Table 2: Deaths due to PM2.5 by Borough - 2017-2019

TimePeriod	Geography	Estimated annual number (age 30+)	Estimated annual rate (age 30+)
2017-2019	Bronx	320.0	37.0
2017-2019	Brooklyn	561.0	35.0
2017-2019	Manhattan	365.0	33.0
2017-2019	Queens	418.0	27.0
2017-2019	Staten Island	134.0	43.0

Time series analysis

In the context of this study, the primary objective is to predict future levels of specific air pollutants such as PM2.5 (particulate matter), O3 (ozone), and the Air Quality Index (AQI). These pollutants are known to have significant impacts on public health, particularly in terms of cardio-respiratory diseases. By analyzing past and present data, time series analysis helps in understanding trends, seasonal patterns, and potential anomalies, which can then inform predictions about future pollution levels. This predictive capability is essential for policymakers and health officials to implement measures aimed at mitigating the adverse health effects of poor air quality.

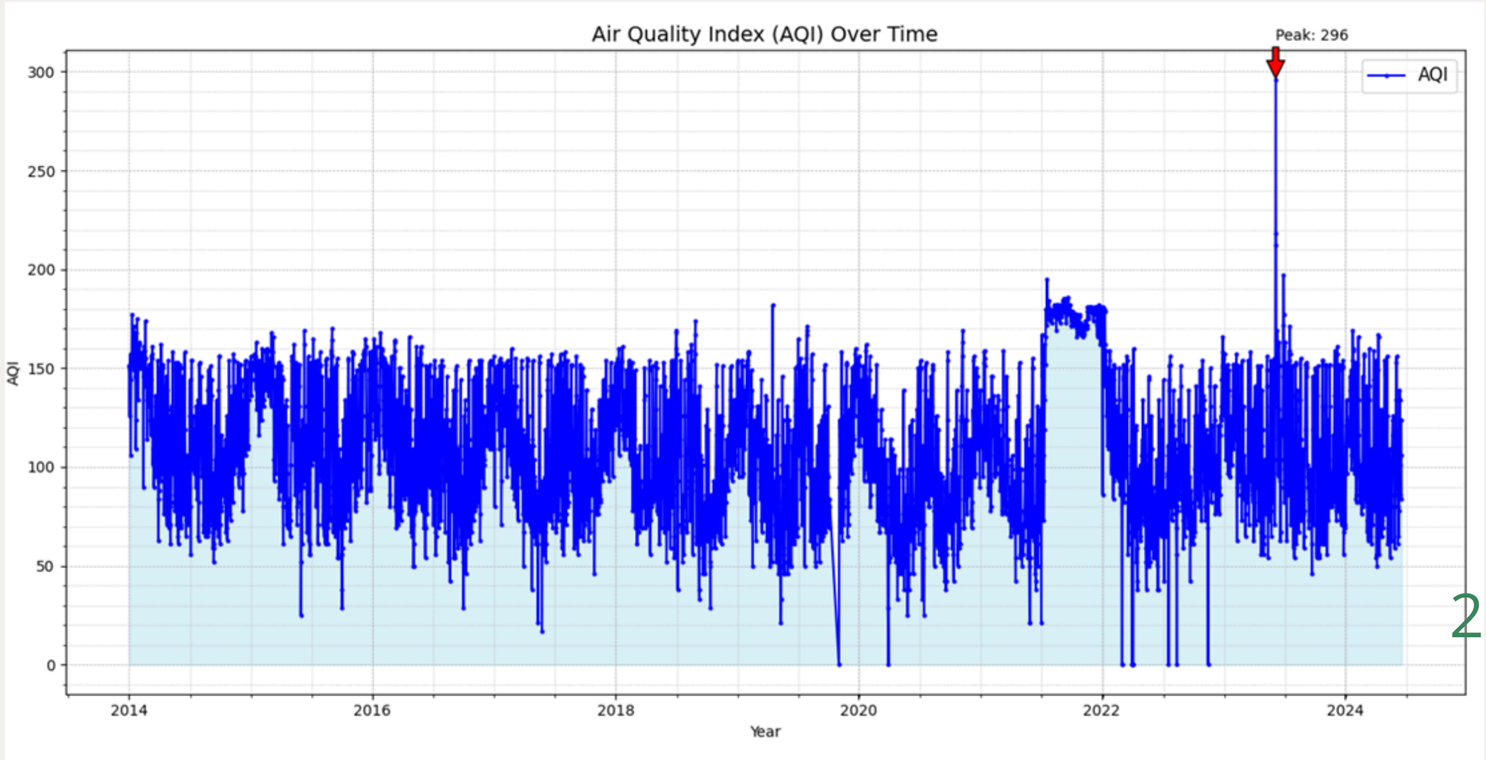
Air Quality Index

The data for this analysis was collected from AQICN.org. The dataset covers a period from 2014 to the present and includes daily measurements of key pollutants such as PM2.5, O3, NO2, and CO. This study does not take into account NO2 an CO.

An essential step in the data preparation process was the calculation of the Air Quality Index (AQI), which provides a comprehensive measure of overall air quality based on the levels of various pollutants.

Table 3: Breakpoints for Calculating AQI from PM2.5 Concentrations

PM2.5 Concentration ($\mu g/m^3$)	AQI Value	AQI Category
0.0 - 12.0	0 - 50	Good
12.1 - 35.4	51 - 100	Moderate
35.5 - 55.4	101 - 150	Unhealthy for sensitive groups
55.5 - 150.4	151 - 200	Unhealthy
150.5 - 250.4	201 - 300	Very unhealthy
250.5 - 350.4	301 - 400	Hazardous
350.5 - 500.4	401 - 500	Hazardous



Data Cleaning and Processing

The data cleaning and processing involved several steps to ensure accuracy and consistency. This included:

- **Converting date formats to a standard structure**
- **Sorting the data chronologically**
- **Removing any duplicate entries**
- **Handling missing values**

Handling missing values was particularly important; missing data points can introduce bias or errors into the analysis. The K-Nearest Neighbors (KNN) Imputer was employed to fill in the missing values. This method uses the values of the five nearest neighbors to estimate and replace missing data points, ensuring a continuous and reliable dataset for subsequent forecasting analysis.

Facebook Prophet

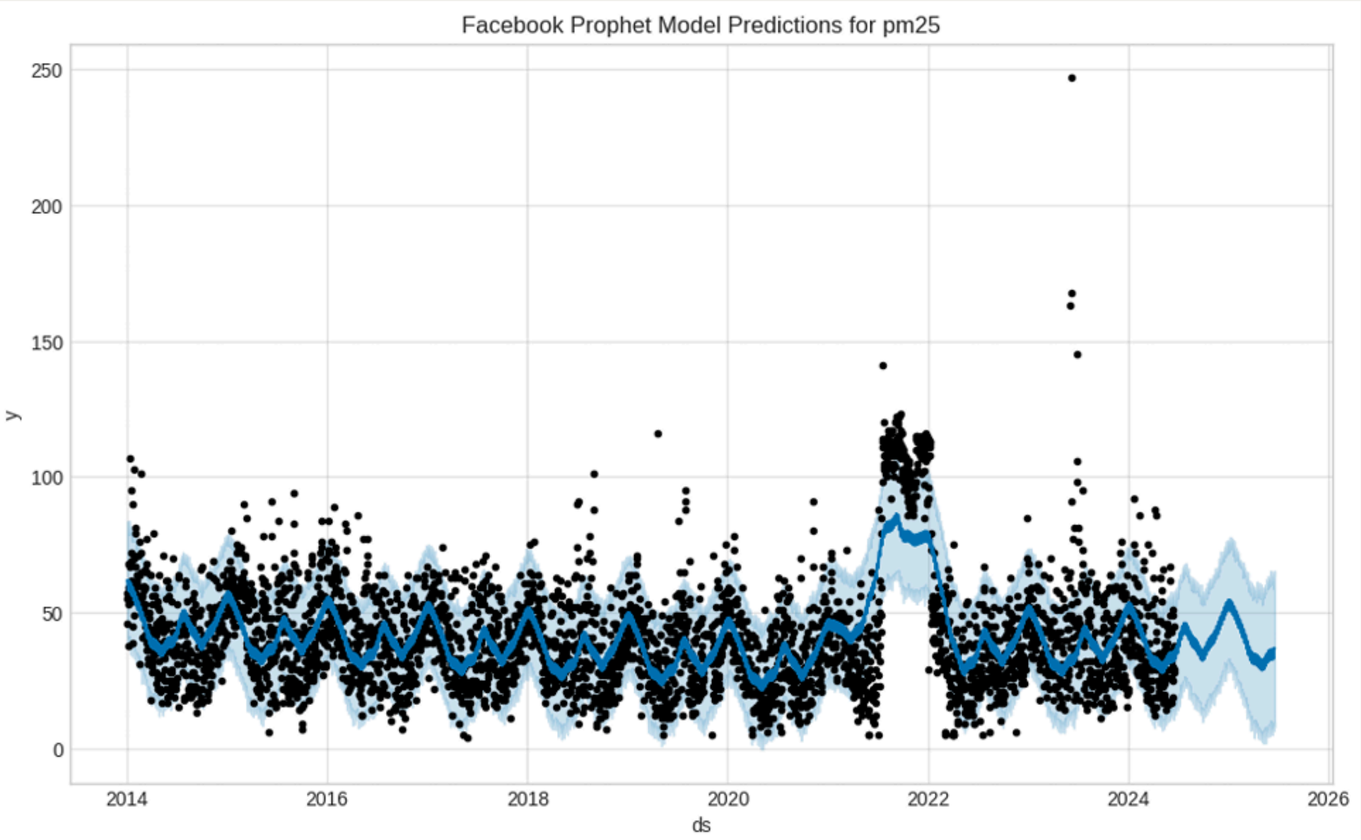
For the time series analysis, Facebook Prophet was utilized as the primary modeling tool. Prophet is a robust forecasting tool developed by Facebook's Core Data Science team, designed to handle time series data with daily observations and missing values, seasonal effects, and trend changes.

Prophet's flexibility and ease of use make it well-suited for forecasting environmental data such as air pollution levels. It decomposes the time series into trend, seasonality, and holidays, providing a comprehensive model that captures the underlying patterns in the data.

So, for each pollutant and AQI, the following steps were taken:

1. Model Training: the dataset was split into training and validation sets. Prophet's parameters were fine-tuned to optimize the model's performance.
2. Forecasting: future values of pollutants and AQI were predicted using the trained model, extending the forecast horizon to anticipate future trends.
3. Visualization: the results were visualized through plots, showing the predicted levels of PM2.5, O3 and AQI over time.

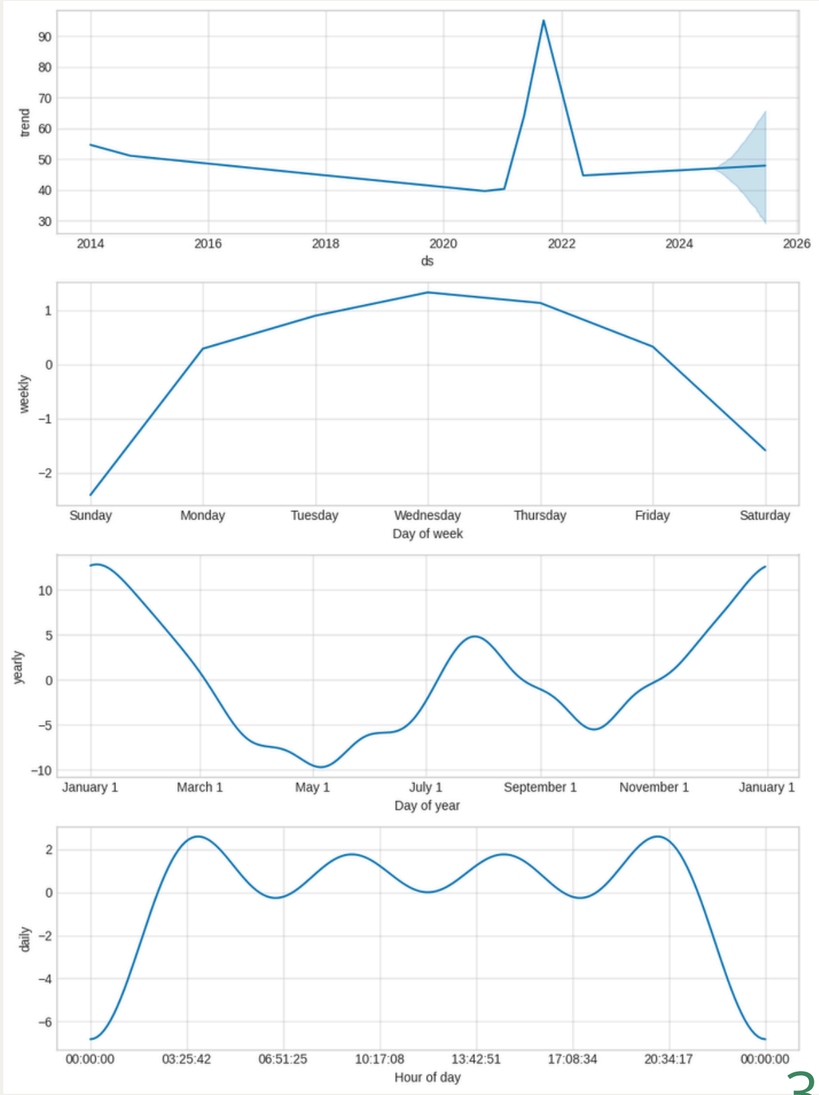
PM2.5 (Fine particulate matter)



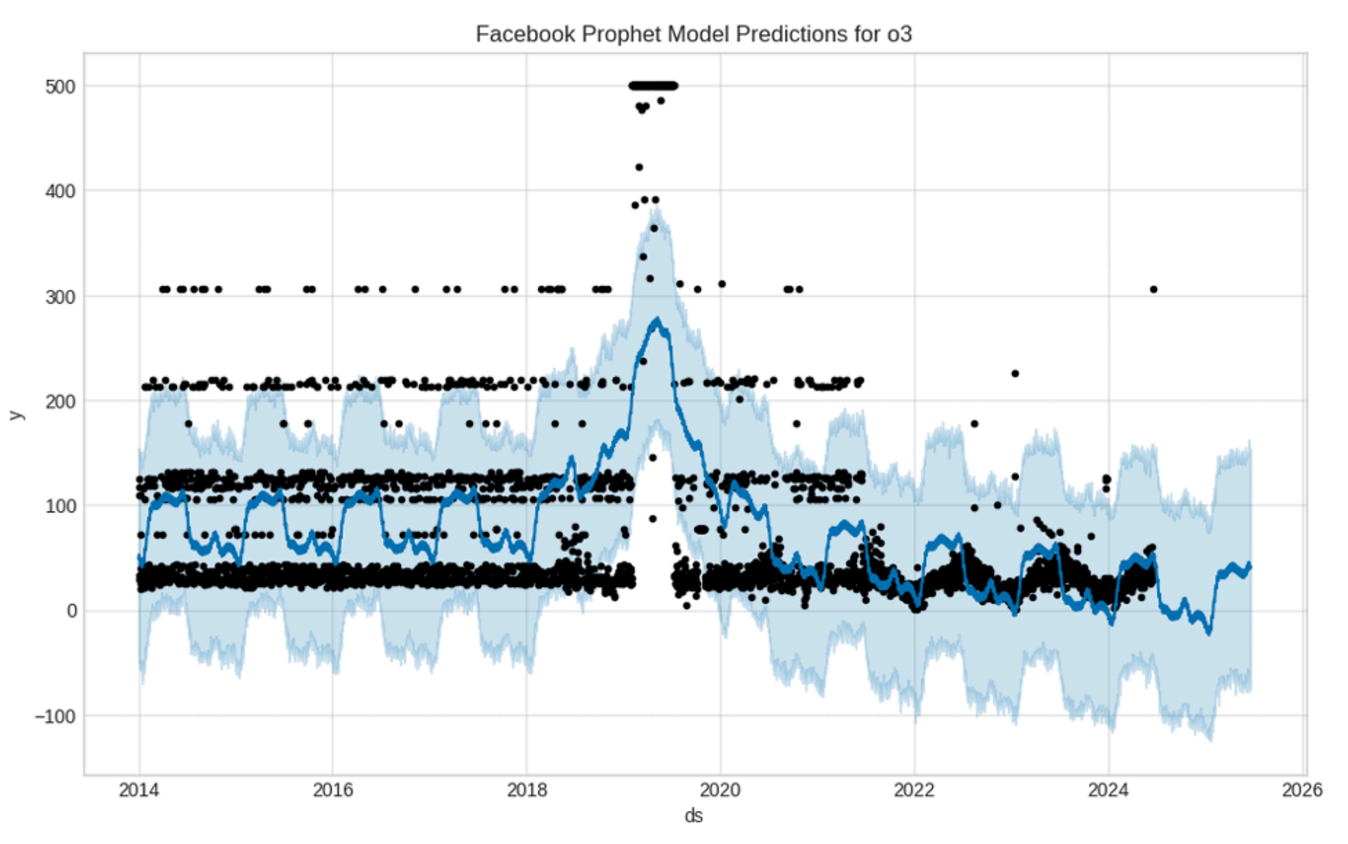
MAE: 13.90

MSE: 318.29

RMSE: 17.84



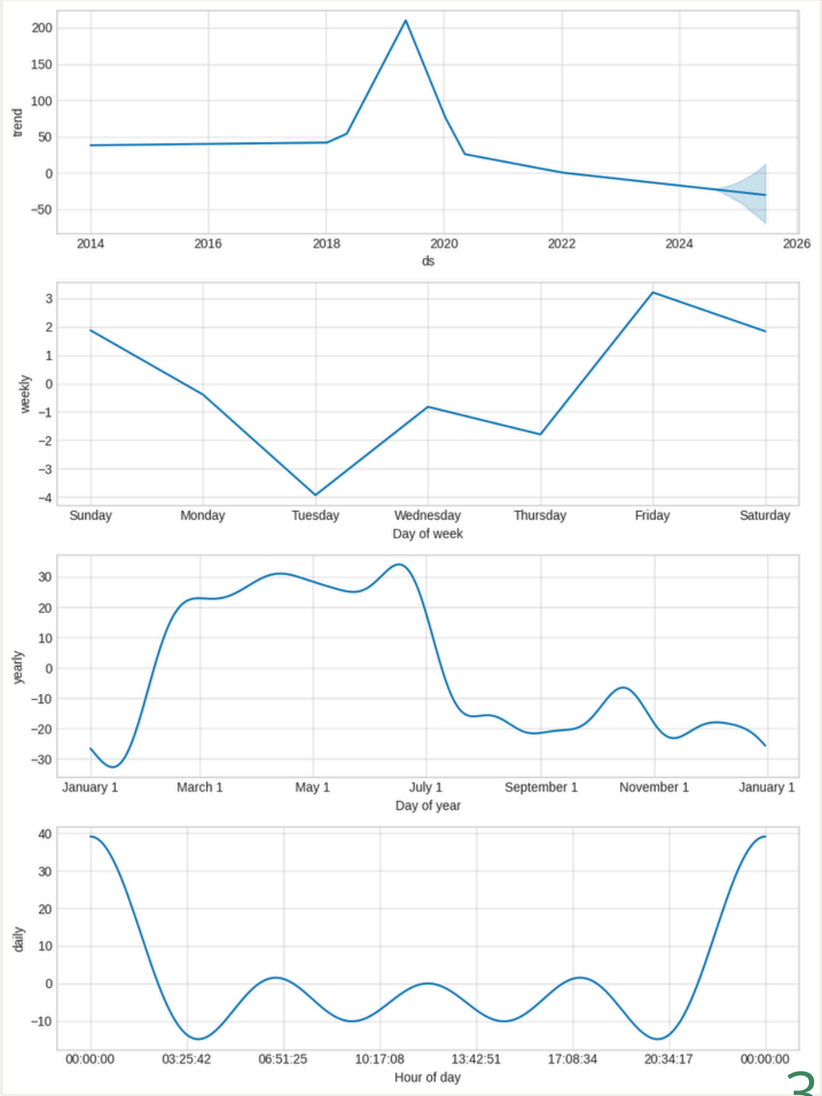
Ozone (O3)



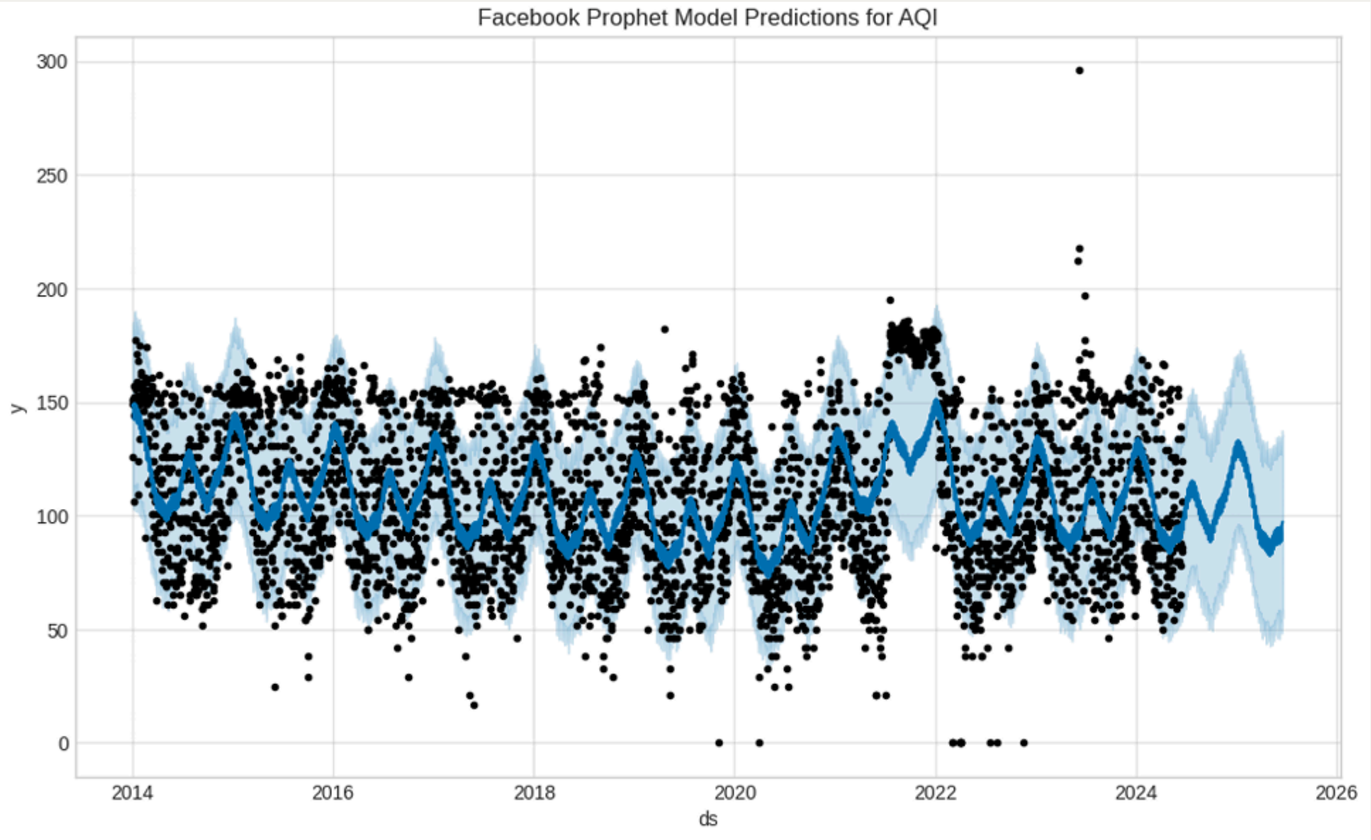
MAE: 23.03

MSE: 1020.51

RMSE: 31.95



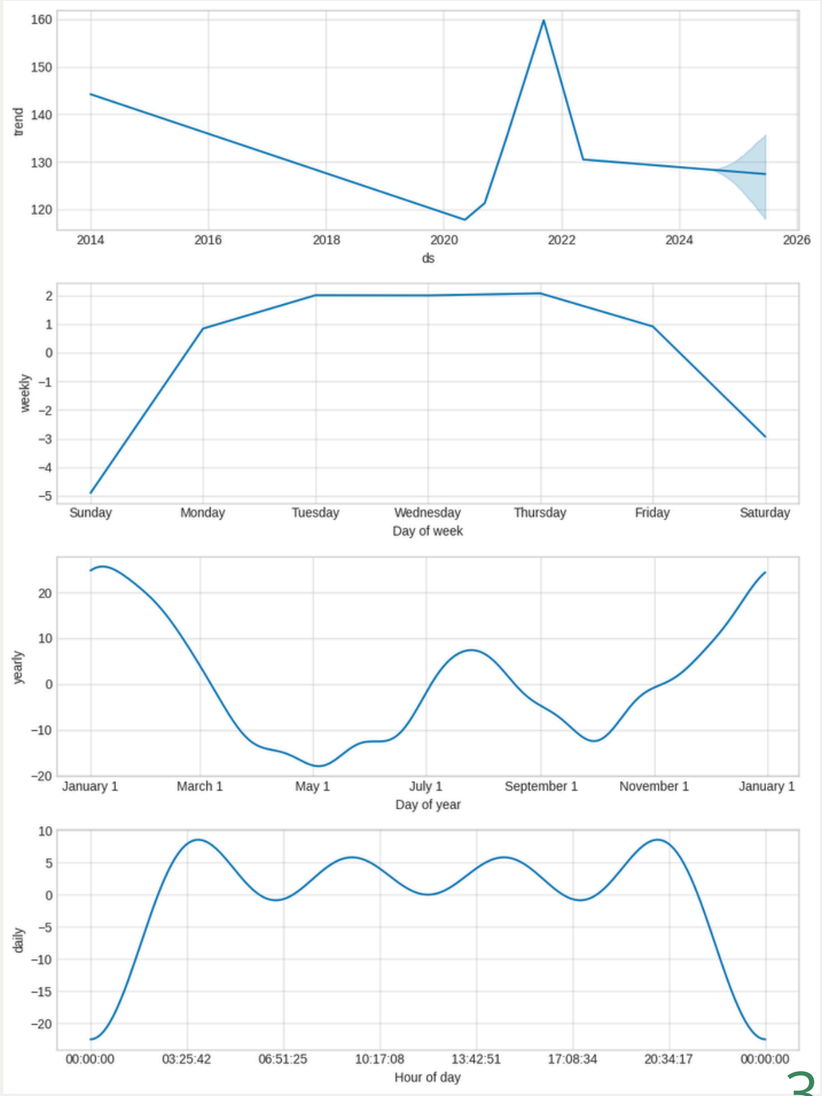
Air Quality Index (AQI)



MAE: 26.93

MSE: 1065.50

RMSE: 32.64



Policy suggestions

Enhance emission standards and regulations by implementing stricter standards for vehicles and industrial sources.

Encourage the adoption of renewable energy sources, such as wind, solar, and hydroelectric power, by providing subsidies and tax incentives to businesses and households.

Urban green spaces should be increased, including parks, green roofs, and urban forests

Improve public transportation and infrastructure for non-motorized transportation such as cycling and walking

Support for vulnerable populations

Believe in community engagement and education

Thank you for your attention