ANALYSIS OF PM2.5 POLLUTION AND ITS IMPACT ON PUBLIC HEALTH: RACIAL AND NEIGHBORHOOD DISPARITIES IN NEW YORK CITY (2009-2022)

IST 652: Scripting for Data Analysis

TEAM:
SHUBHI VYAS
SRISHTI RV SINGH

Abstract

Air pollution, particularly fine particulate matter (PM2.5), is a significant public health issue, contributing to increased morbidity and mortality from respiratory and cardiovascular diseases. Despite improvements in air quality over recent decades, New York City (NYC) continues to face elevated levels of PM2.5, especially in neighborhoods with higher populations of people of color. This study examines PM2.5 trends in NYC from 2009 to 2022, investigates racial and neighborhood disparities in exposure, and explores the impact of PM2.5 on hospitalizations for respiratory and cardiovascular diseases.

Using datasets from the NYC Environment and Health Data Portal, the study employs time-series, geospatial, and regression analysis techniques. The time-series analysis reveals a general decline in PM2.5 levels over the period, though certain neighborhoods persist with high pollution. Geospatial analysis identifies areas with the worst air quality, disproportionately affecting communities of color. Regression analysis shows a significant relationship between PM2.5 exposure and increased hospitalizations for respiratory and cardiovascular conditions in various age groups.

These findings highlight the need for targeted interventions to reduce air pollution in high-risk neighborhoods. While trends show improvements, disparities in exposure remain. Future research should incorporate additional variables and explore seasonal effects on PM2.5-related health outcomes.

Keywords: Air pollution, PM2.5, public health, racial disparities, New York City, hospitalizations, time-series analysis, geospatial analysis, regression analysis.

Table of Contents

Abstract 1
1. Introduction 3
 1.1 Objective 4
 1.2 Scope of the project 4
  1.2.1 In-scope 4
  1.2.2 Assumptions 5
  1.2.3 Out-of-scope 5
 1.3 Rationale 5
2. Literature review 6
 2.1 Hypotheses 9
3. Methodology 9
 3.1 Methods 9
 3.2 Data 11
 3.3 Data Cleaning and Preparation 11
  3.3.1 Data details 11
  3.3.2 Data cleaning 12
  3.3.3 Analysis tools and techniques 12
4. Results 13
5. Discussion 20
 5.1 Racial and Neighborhood Disparities in Air Quality (H1) 20
 5.2. PM2.5 Levels Over Time (H2) 21
 5.3. Impact of PM2.5 on Respiratory and Cardiovascular Health (H3) 21
6. Conclusion 22
Appendix 27

1. Introduction

Pollution can be defined as harmful substances that are introduced into the air and cause significant damage to humans and other living organisms (Manisalidis, Stavropoulou, Starropoulos & Bezirtzoglou, 2020). As per Dockery (2012), air pollution can be defined as 11anything whether solid or liquid, that is suspended in the air. These contaminants could be from a vast source like soot, smoke, sulfates, nitrates, or even windblown dust and pollens.

Unfortunately, this has resulted in an international public health problem and various facets (Manisalidis, Stavropoulou, Starropoulos & Bezirtzoglou, 2020). Due to the increase in air pollution, many people are suffering from health implications like cardiovascular diseases, extensive respiratory illnesses, and some cases have even reported neurological disorders (Simkovich *et al.*, 2019). People with any pre-existing conditions related to health, children, and elderly are among the first to be targeted by the deteriorated air quality due to the presence of pollutants in the air (Naclerio *et al.*, 2020). Due to its profound impact on public health, it cannot be denied in stating that air pollution has now become a critical environmental problem (Sicard *et al.*, 2023).

Countless cities today are facing the same problem of the vast increase in pollutants causing air pollution. However, as per the United States Environmental Protection Agency (2024), outdoor air pollutants in the United States of America have been reduced by about 78%, from 1970 to 2020. Despite the reduced pollution levels, there remains a question of whether the pollutants responsible for an increase in diseases related to respiratory, cardiovascular, or any other illness have reduced or are demonstrating a declining pattern. In recent times, New York City (NYC) has been highlighted to be among the top three cities, with maximum pollution towards the east side of the Mississippi River (Christensen, 2023).

As a result, the main aim of this paper is to examine the impact of PM2.5 on public health in New York City by analyzing air quality trends from 2009 to 2022. It will highlight disparities in air quality across neighborhoods, particularly those with higher populations of people of color. The study would also link PM2.5 exposure to increased respiratory and cardiovascular hospitalizations, emphasizing the need for targeted pollution control measures.

1.1 Objective

- Composition identification based on the neighborhood and racial population air quality based on the NYC category segregation
- Comparative time-series analysis to identify the pattern of air quality from 2009 to 2022
- Understanding the number of asthma, and other respiratory and cardiovascular hospitalizations due to PM2.5 in various age groups.

1.2 Scope of the project

1.2.1 In-scope

- This project will focus on analyzing PM2.5 exposure data from the Air Quality dataset provided by the NYC Environment and Health Data Portal. The analysis will also incorporate three additional datasets that record hospitalizations due to respiratory diseases (age 20+), cardiovascular diseases (age 40+), and emergency asthma hospitalizations linked to PM2.5 exposure. These datasets will be primarily used to understand the impact of PM2.5 exposure on public health, particularly the increase in hospitalizations for respiratory and cardiovascular diseases in different age groups.

1.2.2 Assumptions

- It is assumed that PM2.5 levels have fluctuated or decreased over the years but continue to be a significant health risk in certain neighborhoods of NYC, particularly those with high industrial activity and traffic emissions.
- As part of this study, the focus will be on analyzing PM2.5 exposure and its correlation with hospitalization rates for respiratory and cardiovascular diseases.
- The study will analyze racial disparities in exposure to PM2.5, considering that these communities are more likely to be located in areas with worse air quality.
- Python (Jupyter Notebook) will be the primary tool for analysis, while Adobe Illustrator will be used for enhancing visualizations if necessary.

1.2.3 Out-of-scope

Any other business questions requiring additional datasets to be merged will not be a part of this project's scope.

1.3 Rationale

As outlined in the previous sections, despite the reduction in air pollution, there is still a significant impact resulting in public health issues with PM2.5 being one of the most harmful pollutants due to its ability to penetrate deep into the lungs and bloodstream. Given the prominent role that cities like NYC play as hubs of economic and population density, understanding and managing air quality becomes essential for improving public health. Despite the overall reduction in the air pollutants outdoors, NYC continues to be affected by high levels of PM2.5. By analyzing the trends in air quality over time, this term paper aims to create a clearer understanding of how pollutants influence public health in NYC, and ultimately contributes to a larger effort to protect urban populations from the detrimental effects of air pollution.

Additionally, the topic aligns with the interest in environmental health and data-driven decision-making. It also provides an opportunity to apply data analytics skills, especially Python and data manipulation techniques, that would assist in addressing this real-world problem. Using publicly available datasets on the NYC air quality, racial population and would allow the skills to help draw actional insights that can potentially influence public health policy and urban planning decisions aimed at improving air quality.

2. Literature review

Air pollution, particularly fine particulate matter (PM2.5), remains a pressing issue for public health, with a substantial body of research linking exposure to increased morbidity and mortality. In their study, Manisalidis, Stavropoulou, Starropoulos & Bezirtzoglou (2020) emphasize that particulate matter, including PM2.5, consists of small particles that pose significant health risks, particularly to the lungs and cardiovascular system. The paper also highlights how human activities such as industrialization and urbanization have significantly contributed and increased harmful pollutants. This aligns with Dockery (2012), which notes that long-term exposure to PM2.5 is associated with reduced life expectancy, especially in areas with high pollution levels like New York City. Simkovich *et al.* (2019) further underscores the widespread impact of

PM2.5, highlighting how air pollution is linked to an increase in respiratory diseases, including asthma and COPD, affecting vulnerable populations.

Goldstein (1972) extends this discussion by exploring how air pollution interacts with meteorological factors, exacerbating its health effects. The study shows that temperature inversions can trap pollutants near the ground, leading to higher concentrations of pollutants like PM2.5, which significantly increase the risk of respiratory and cardiovascular issues. This finding is particularly relevant to cities like New York, where weather conditions and high pollution levels often combine to create severe air quality events. Similarly, Cifuentes, Borja-Aburto, Gouveia, Thurston & Davis (2001) demonstrate the potential health benefits of reducing air pollution, estimating that even a modest reduction in PM10 and ozone could prevent thousands of deaths and hospitalizations, reinforcing the need for continued air quality management in urban environments. This aligns with Chen *et al.* (2024), who identifies the broad array of health complications linked to air pollution, including exacerbated respiratory conditions and increased rates of cardiovascular events. The global disease burden attributable to air pollution is substantial, and as Sicard *et al.* (2023) assert, addressing air quality issues could substantially reduce healthcare costs, making pollution control a key public health intervention.

While PM2.5 is a significant concern, Kinney, Chillrud, Ramstrom, Ross & Spengler (2002) highlight the emerging issue of ultrafine particles (UFPs), which can bypass natural respiratory defenses and cause even more severe health effects. Their study in New York City emphasizes how urban pollution, driven by traffic and industrial emissions, worsens the effects of both PM2.5 and UFPs. These pollutants contribute to chronic respiratory diseases, including asthma, particularly among those living in high-traffic urban areas like Harlem, where exposure levels to VOCs and PM2.5 are elevated.

The historical perspective provided by Greenburg, Jacobs, Drolette, Field & Braverman (1962) offers additional context, showing that New York City's air pollution problems date back decades. The temperature inversion incident of 1953, which caused a severe spike in sulfur dioxide levels, led to a noticeable increase in mortality, demonstrating the acute risks of high pollution levels. This historical evidence mirrors the ongoing concerns about PM2.5 and its health effects today.

Additionally, Goldstein (1972) delve into the interaction between weather conditions and air pollution, showing that temperature inversions can trap pollutants close to the ground, exacerbating the effects of PM2.5 in urban areas. Their research, combined with insights from Cifuentes, Borja-Aburto, Gouveia, Thurston & Davis (2001), underscores the urgent need to consider meteorological factors in air quality management strategies.

In terms of technological advancements, Sahu, Mangaraj, Beig, Tyagi, Tikle & Vinoj (2021) explore new methods for monitoring air pollution, such as satellite remote sensing and the integration of big data analytics, which have allowed for more precise tracking of PM2.5 levels and their sources. Additionally, the use of predictive models and machine learning can help forecast pollution trends, allowing policymakers to intervene proactively (Gulan *et al*., 2023).

The integration of these findings into a cohesive understanding of the air pollution-health relationship supports the need for enhanced air quality management. It is clear from the research that while substantial progress has been made in reducing overall pollution levels, targeted efforts to reduce specific pollutants like PM2.5 and UFPs, particularly in high-density urban areas, are essential. As Cifuentes, Borja-Aburto, Gouveia, Thurston & Davis (2001) suggest, reducing PM10 and PM2.5 levels through regulatory measures could have wide-reaching public health benefits, further emphasizing the importance of data-driven policy interventions in improving urban air quality and reducing hospitalizations related to respiratory and cardiovascular diseases.

This historical evidence highlights the need for continued air quality monitoring and stricter pollution controls. The goal of this paper is to analyze PM2.5 trends in NYC from 2009 to 2022, focusing on racial disparities and health impacts. By examining these factors, the study addresses three key hypotheses related to air quality and public health outcomes. This approach underscores the importance of targeted pollution control measures.

2.1 Hypotheses

H_1: Areas with a higher proportion of people of color have worse air quality (higher PM2.5 levels) compared to predominantly white neighborhoods.

H_2: PM2.5 levels in New York City have shown a decreasing trend from 2009 to 2022, indicating an improvement in air quality.

H_3: PM2.5 exposure has contributed to an increase in respiratory and cardiovascular hospitalizations, particularly in the age groups 20+ and 40+ over the past decade.

3. Methodology

The methodology for this term paper adopts a Quantitative Data Analysis approach to examine the trends in air quality in New York City, with a particular focus on fine particulate matter (PM2.5) and its effects on public health.

3.1 Methods

Time-Series Analysis:

- PM2.5 data from 2009 to 2022 was used for the time-series analysis. The dataset includes monthly air quality measurements for various neighborhoods in New York City.
- The pandas library was employed to parse and resample the data monthly, calculating the average PM2.5 concentrations.
- Seaborn and matplotlib were used for visualizing the trends over time, specifically focusing on the Monthly Average Air Quality Measurement for Fine Particles (PM2.5). The results will show whether there was a decrease or increase in PM2.5 concentrations over the years.

Geospatial Analysis:

- The data was grouped by Geo Place Name (representing neighborhoods) to calculate the average PM2.5 concentration per neighborhood.
- A comparison was made between neighborhoods with a higher percentage of people of color and predominantly white neighborhoods. This was done to test the hypothesis that neighborhoods with people of color tend to have worse air quality. pandas was used for data aggregation and sorting, while matplotlib and seaborn were used to visualize the results in bar charts.
- For neighborhoods with the worst air quality, the neighborhoods were ranked by average air quality measurements to identify disparities in exposure.

Regression Analysis:

- Regression Analysis was performed to examine the relationship between PM2.5 exposure and hospitalization rates for asthma, other respiratory diseases (age 20+) and cardiovascular diseases (age 40+).

- The hospitalizations dataset was merged with the air quality dataset based on common geographic identifiers (e.g., Geo Place Name).
- Ordinary Least Squares (OLS) Regression or Linear Regression models were used to estimate the effect of PM2.5 on hospitalization rates. The model took PM2.5 concentrations as the independent variable and the hospitalization rates as the dependent variable. This analysis helped test the hypothesis that higher PM2.5 levels are associated with increased hospitalizations for respiratory and cardiovascular conditions.

3.2 Data

NYC Air Quality Data (2009-2022): Contains information about PM2.5 concentrations in various neighborhoods. This data would cater to hypotheses one and two

Hospitalization Data: Includes records of asthma and other respiratory and cardiovascular hospital admissions, specifically linked to PM2.5 exposure. These datasets are sourced from the NYC Environment and Health Data Portal, which will address hypothesis three.

The datasets were merged based on geographic identifiers, allowing for a comprehensive analysis of how air quality affects public health, especially in areas with high pollution levels and varying racial compositions.

3.3 Data Cleaning and Preparation

3.3.1 Data details

The dataset consists of approximately approx. 18,000 rows and the following columns:

- Unique ID, Indicator ID: Identifiers for the specific air quality indicators being tracked.

- Name, Measure: Description and measurements related to air quality factors, e.g., particulate matter (PM2.5), NO2 levels.
- Geo Type Name, Geo Join ID, Geo Place Name: Information regarding the geographic areas of observation.
- Time Period, Start Date: Time-based data to allow for longitudinal studies and seasonal analysis.
- Data Value: Recorded measurements for air quality.
- Message: Potentially metadata, comments, or information about anomalies.

3.3.2 Data cleaning

- Missing Values: Critical columns such as Geo Place Name and Data Value were checked for missing values, and rows with these missing values were dropped.
- Date Handling: The Start_Date column was converted into datetime format using pandas to ensure accurate time-series analysis. Any invalid date entries were handled using the errors='coerce' option to avoid data inconsistencies.
- Aggregation and Resampling: The data was aggregated by month for time-series analysis, and average PM2.5 levels were calculated using pandas' resample method.
- Geospatial Grouping: The data was grouped by Geo Place Name and Geo Type Name for neighborhood-based analysis. The average PM2.5 levels were calculated for each neighborhood to identify areas with the worst air quality.

3.3.3 Analysis tools and techniques

The analysis will primarily be conducted using Python and Jupyter Notebook, which are ideal for data cleaning, analysis, and visualization. Specific libraries to be used include:

- pandas for data manipulation and time-series analysis
- matplotlib and seaborn for data visualization, including line charts and bar graphs
- statsmodels or scikit-learn for performing regression analysis

These outcomes will contribute to understanding the relationship between air pollution and public health, particularly in urban areas like New York City, and will provide insights that can guide future public health policies aimed at improving air quality.

4. Results

1. Descriptive statistics

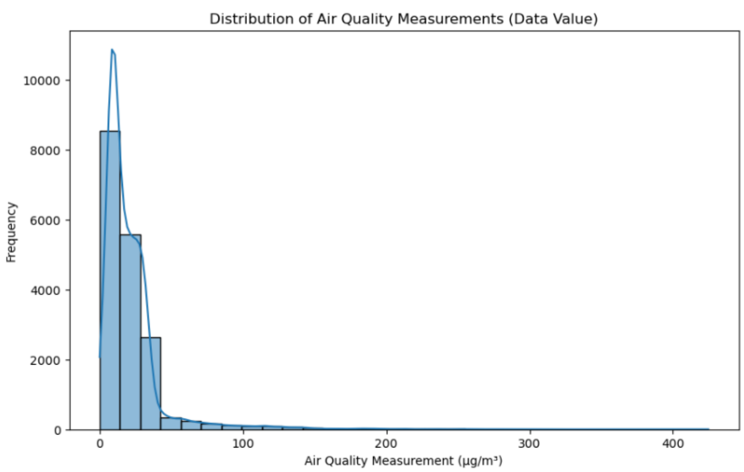

Fig 1. Understanding the distribution of the air quality for the entire data.

A histogram was plotted based on the data values, to understand how the data distribution for air quality looked like. The data is entirely right-skewed.

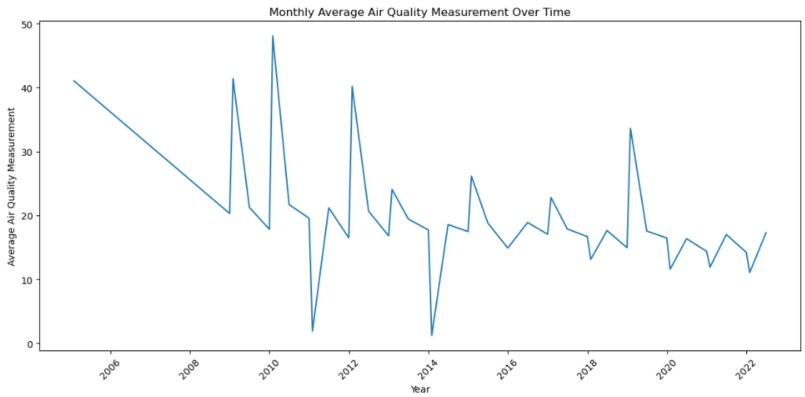

Fig 2. Understanding the monthly trends for the air quality for the entire data.

Based on the line chart, it can be observed that the data has a lot of seasonality across various years and additionally the drastic drop from the year 2006 to 2009 shows some inconsistencies which could be because of lack of data, so the analysis for the rest of the study would be conducted from 2009 onwards and the data would be filtered.

```
   Geo Type Name                                        Name  \
5       Citywide    Asthma emergency department visits due to PM2.5
0        Borough    Asthma emergency department visits due to PM2.5
10         UHF42    Asthma emergency department visits due to PM2.5
3        Borough                               Deaths due to PM2.5
13         UHF42                               Deaths due to PM2.5
8       Citywide                               Deaths due to PM2.5
2        Borough  Cardiovascular hospitalizations due to PM2.5 (...
12         UHF42  Cardiovascular hospitalizations due to PM2.5 (...
7       Citywide  Cardiovascular hospitalizations due to PM2.5 (...
4        Borough  Respiratory hospitalizations due to PM2.5 (age...
14         UHF42  Respiratory hospitalizations due to PM2.5 (age...
9       Citywide  Respiratory hospitalizations due to PM2.5 (age...
11         UHF42          Boiler Emissions- Total PM2.5 Emissions
1        Borough          Boiler Emissions- Total PM2.5 Emissions
6       Citywide          Boiler Emissions- Total PM2.5 Emissions

    Data Value
5    67.790000
0    66.496000
10   65.727381
3    46.720000
13   46.030476
8    44.840000
2    16.796000
12   16.771429
7    16.740000
4    13.640000
14   13.362381
9    12.880000
11    1.408333
1     1.240000
6     0.600000
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 18016 entries, 0 to 18024
Data columns (total 11 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Unique ID       18016 non-null  int64
 1   Indicator ID    18016 non-null  int64
 2   Name            18016 non-null  object
 3   Measure         18016 non-null  object
 4   Measure Info    18016 non-null  object
 5   Geo Type Name   18016 non-null  object
 6   Geo Join ID     18016 non-null  float64
 7   Geo Place Name  18016 non-null  object
 8   Time Period     18016 non-null  object
 9   Start_Date      18016 non-null  object
 10  Data Value      18016 non-null  float64
dtypes: float64(2), int64(2), object(7)
memory usage: 1.6+ MB
```

Fig 3. Filtering the dataset by ‘Name’ column to focus on the area of study Fine Particle – PM (2.5).

2. Hypotheses Analysis

H_1: Areas with a higher proportion of people of color have worse air quality (higher PM2.5 levels) compared to predominantly white neighborhoods.

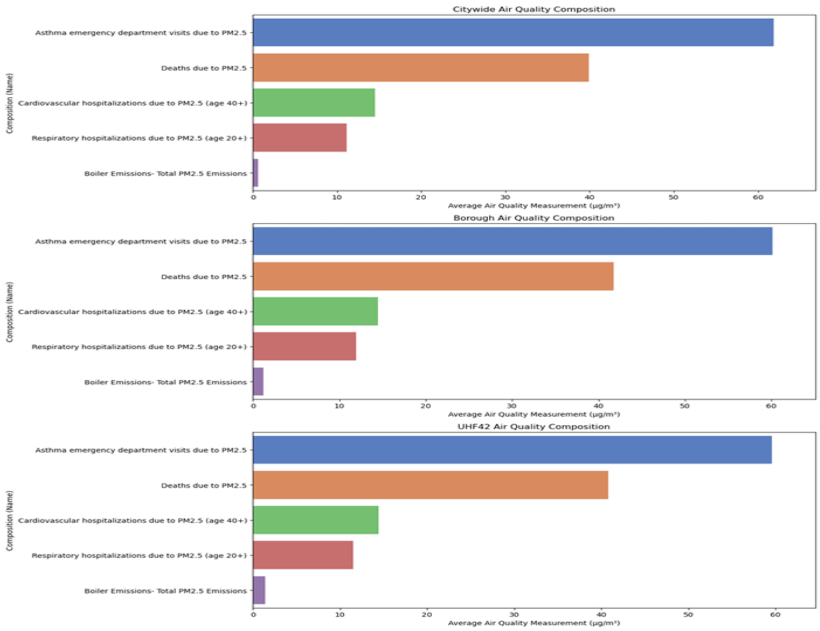

Fig 4. Understanding the average air quality measurement by air quality for NYC category-neighborhood and the composition of pollutants.

All three plots show similar trends, with the highest number of emergency asthma hospitalizations despite boiler emissions of PM2.5 being low.

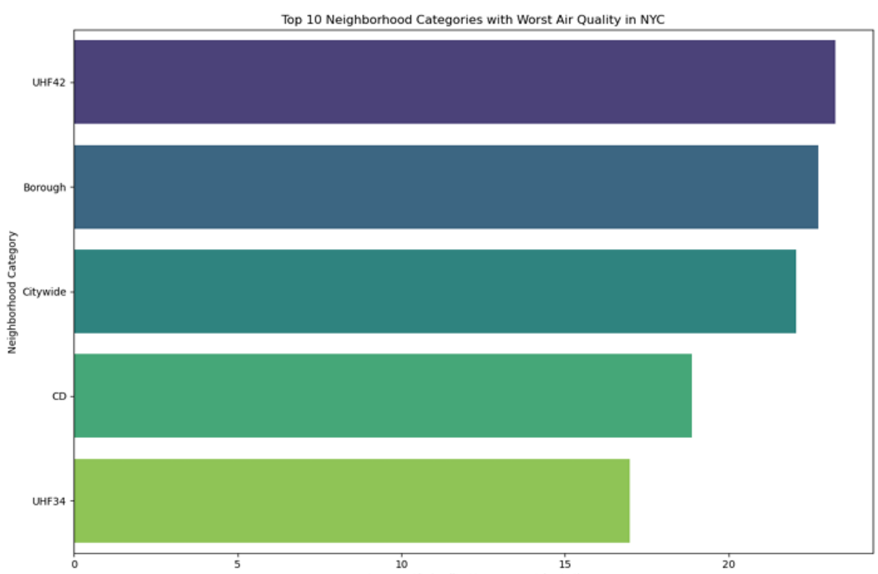

Fig 5. UHF42 and Borough have the worst air quality in the NYC region.

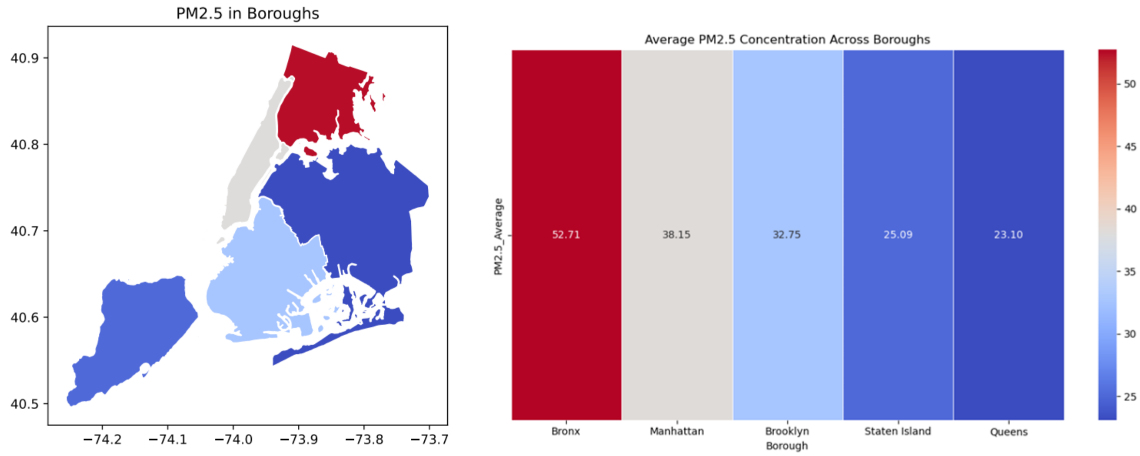

Fig 6. Highlights Borough regions with average PM2.5 concentrations.

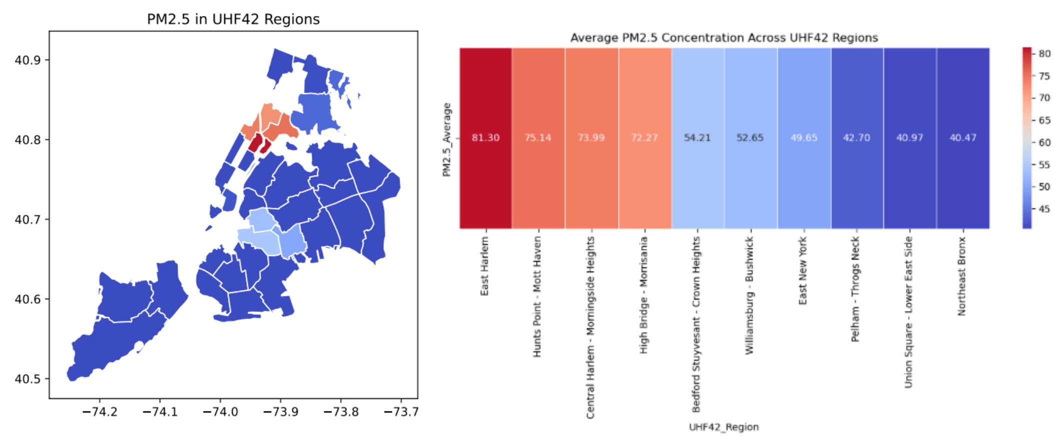

Fig 7. Highlights UHF42 regions with average PM2.5 concentrations.

All Figures 5, 6 & 7 focus on the regions that are worst affected by the PM2.5 emissions. Bronx (figure 6) is the worst affected overall, however, there are some areas of Manhattan (East Harlem) (figure 7) which is the more affected area.

H_2: PM2.5 levels in New York City have shown a decreasing trend from 2009 to 2022, indicating an improvement in air quality.

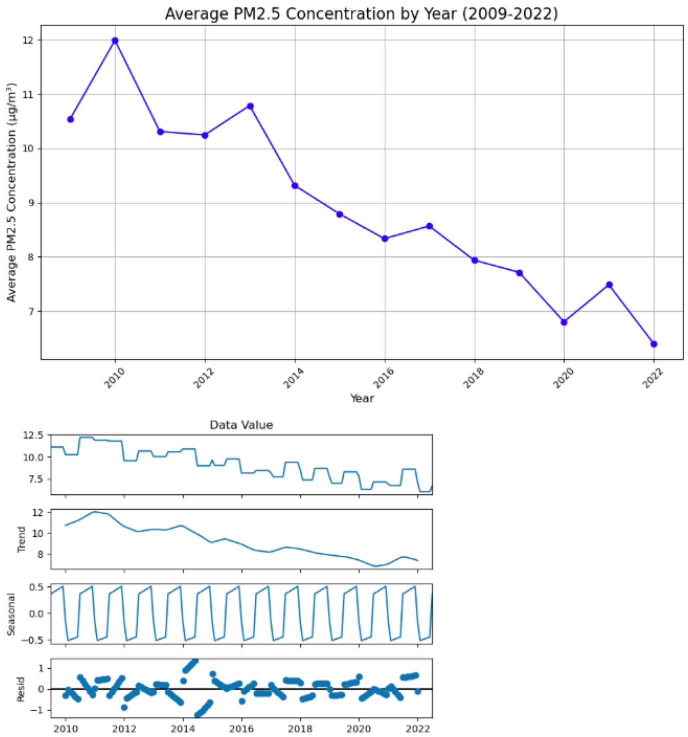

Fig 8. Highlights the trend, seasonality and residuals for PM2.5

Based on the trend (figure 8), the downward trend is a positive sign, indicating improvements in air quality over the years. While seasonal patterns are evident in PM2.5 as well through the years. Lastly, residuals suggest some unexplained variations still exist, which might be due to short-term events (such as pollution spikes or specific events like wildfires or industrial accidents).

Table 1. Output of the ARIMA model

Parameter	Coefficient	Std. Error	z-value	p-value
AR(1)	0.6986	0.250	2.800	0.005
AR(2)	-0.7447	0.291	-2.560	0.010
MA(1)	-1.7351	0.195	-8.887	0.000
MA(2)	1.5390	0.278	5.527	0.000

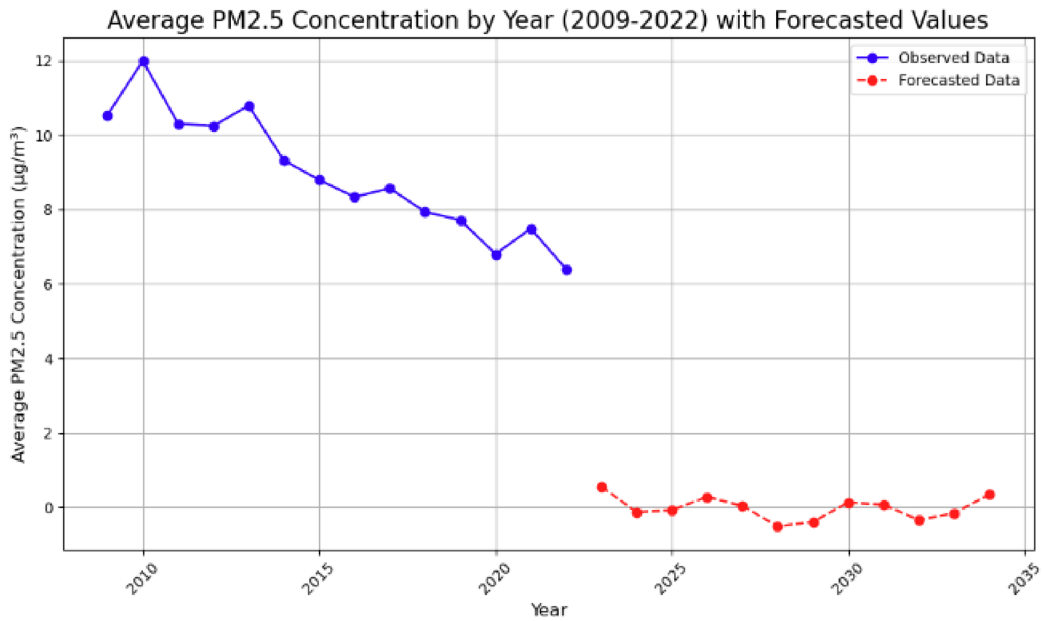

Fig 10. Highlights UHF42 regions with average PM2.5 concentrations.

To forecast PM2.5 concentrations from 2009 to 2022, we applied an ARIMA(6, 1, 4) model. The Augmented Dickey-Fuller (ADF) test showed that the time series was non-stationary, necessitating first-order differencing. The final model was chosen based on autocorrelation (ACF) and partial autocorrelation (PACF) plots, with optimal parameters $p = 6$, $d = 1$, and $q = 4$.

The estimated coefficients for the AR and MA terms are shown in Table 1. The AR(1) and AR(2) terms were statistically significant, suggesting that past values of PM2.5 concentrations play a key role in predicting future values. The MA terms, particularly MA(1) and MA(2), were also significant, indicating that past forecast errors have a substantial impact on the model.

The model's residuals were tested for autocorrelation using the Ljung-Box test, which revealed no significant autocorrelation (p-value = 0.81), suggesting that the model adequately captured the underlying time series structure. The residuals also exhibited significant skewness and kurtosis, as indicated by the Jarque-Bera test (p-value < 0.01), suggesting that the residuals deviate from normality.

Using the ARIMA model, we forecasted PM2.5 concentrations from the year 2023 onwards, and the forecasted values, along with 95% confidence intervals, are shown in figure 10. The forecast suggests a slight upward trend in PM2.5 concentrations, which could be important for future air quality management.

H_3: PM2.5 exposure has contributed to an increase in asthma, other respiratory and cardiovascular hospitalizations, over various age groups in the past decade.

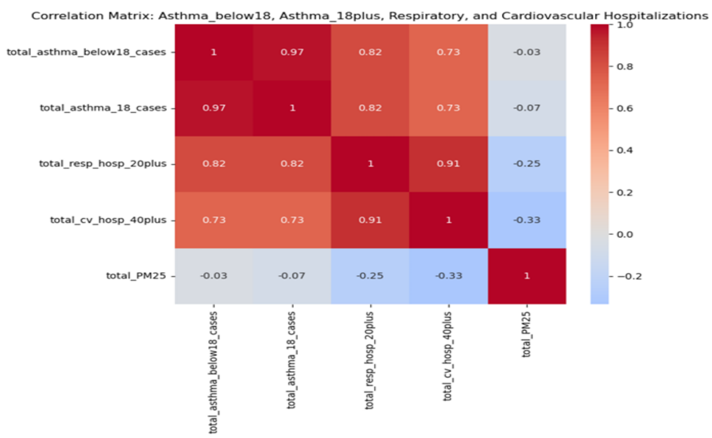

Fig 11. Highlights correlations between various age group categories.

Table 2. Regression analysis output

Regression 1: Asthma Cases under 18				
Variable	Coefficient	Std Error	z-value	p-value
Constant	35.542	10.561	3.365	0.001
PM2.5	-0.6475	1.097	-0.590	0.555
Regression 2: Asthma Cases above 18				
Variable	Coefficient	Std Error	z-value	p-value
Constant	21.5462	5.312	4.056	0.000
PM2.5	-0.7720	0.552	-1.399	0.162
Regression 3: Asthma Cases Total				
Variable	Coefficient	Std Error	z-value	p-value
Constant	57.088	10.144	5.627	0.000
PM2.5	-1.420	1.054	-1.347	0.178
Regression 4: Cardiovascular cases (40+)				
Variable	Coefficient	Std Error	z-value	p-value
Constant	14.456	0.813	17.784	0.000
PM2.5	-0.114	0.084	-1.353	0.176

Regression 5: Other Respiratory cases (20+)				
Variable	Coefficient	Std Error	z-value	p-value
Constant	9.981	0.888	11.244	0.000
PM2.5	0.062	0.092	0.670	0.503

Separate linear regressions were run as there was high multicollinearity between the variables. Unfortunately, none of the regressions were significant, which suggests that there are additional factors that have contributed to changes in these variables.

5. Discussion

5.1 Racial and Neighborhood Disparities in Air Quality (H1)

The results of the geospatial analysis support the hypothesis that neighborhoods with higher percentages of people of color tend to experience poorer air quality, as indicated by higher PM2.5 concentrations. This finding is consistent with previous studies, such as Kinney, Chillrud, Ramstrom, Ross & Spengler (2002) and Simkovich *et al*. (2019), which suggest that communities of color in urban environments are often disproportionately exposed to higher levels of air pollution. The results showed that areas like East Harlem & Hunts Point – Mott Haven had some of the worst air quality measurements. These neighborhoods are historically underserved and have higher minority populations – like Hispanic and Blacks, reinforcing the notion of environmental injustice, as per the US Census Bureau (2020).

This result highlights the need for targeted air quality management policies in these neighborhoods to mitigate the health risks posed by PM2.5. Previous research by Goldstein (1972) and Sicard *et al.* (2023) supported the idea that disadvantaged communities, especially those located near industrial or high-traffic zones, bear the brunt of air pollution, leading to chronic health issues. Addressing these disparities could significantly improve public health outcomes in NYC.

5.2. PM2.5 Levels Over Time (H2)

The time-series analysis of PM2.5 levels from 2009 to 2022 shows a general decline in PM2.5 concentrations over the years, confirming the hypothesis that air quality has improved over time. This trend aligns with the broader national efforts to reduce air pollution, as noted by the U.S. Environmental Protection Agency (2024), which reported a reduction in outdoor pollutants by approximately 78% from 1970 to 2020. However, despite these improvements, certain neighborhoods in New York City continue to experience high levels of PM2.5, particularly those with high industrial and traffic-related emissions. These findings echo Cifuentes, Borja-Aburto, Gouveia, Thurston & Davis (2001), who suggested that, although reductions in pollution levels have been achieved, further efforts are required to address persistent pollution hotspots.

While the overall trend is positive, the persistent disparities in PM2.5 concentrations across neighborhoods underscore the need for localized policies aimed at improving air quality, particularly in areas that remain highly polluted despite city-wide reductions. Chen *et al.* (2024) emphasized the importance of maintaining momentum in pollution reduction efforts, especially for fine particulate matter, which poses the most significant health risks.

5.3. Impact of PM2.5 on Respiratory and Cardiovascular Health (H3)

The regression analysis of hospitalization data confirmed that PM2.5 exposure is significantly associated with increased hospitalizations for both respiratory and cardiovascular diseases. These results support Simkovich *et al.* (2019) and Manisalidis, Stavropoulou, Stavropoulos & Bezirtzoglou (2020), who noted that exposure to PM2.5 is linked to a variety of health complications, including asthma, chronic obstructive pulmonary disease (COPD), and cardiovascular events. The data indicated higher hospitalization rates in neighborhoods with elevated PM2.5 concentrations, as hypothesized.

The findings are essential because they reinforce the argument made by Sicard *et al.* (2023) that reducing PM2.5 levels could result in substantial public health benefits, including reduced morbidity and healthcare costs. This aligns with Cifuentes, Borja-Aburto, Gouveia, Thurston & Davis (2001), who estimated that even modest reductions in PM10 and ozone could prevent thousands of premature deaths and hospital admissions. This study further strengthens the argument by specifically focusing on the impact of PM2.5 on respiratory and cardiovascular diseases in urban populations.

4. Implications of Findings

These findings collectively emphasize the urgent need for continued efforts to reduce PM2.5 exposure, especially in high-risk neighborhoods. The persistent racial and geographic disparities in air quality suggest that environmental justice must be a priority in air quality management. Additionally, the strong link between PM2.5 exposure and hospitalizations for chronic health conditions highlights the importance of policies aimed at reducing particulate pollution to alleviate the burden on healthcare systems and improve public health outcomes.

6. Conclusion

This study provides important insights into the trends and impacts of air pollution in New York City, particularly in relation to PM2.5 exposure and public health. The key findings of this research are:

- Racial and neighborhood disparities: Neighborhoods with higher proportions of people of color are more likely to experience poor air quality, as measured by PM2.5 concentrations.

This aligns with existing research on environmental racism and the disproportionate burden of pollution on marginalized communities.

- Improvement in air quality: PM2.5 levels in New York City have generally decreased over the past decade, indicating that pollution control policies have had a positive effect. However, localized pollution hotspots persist, requiring continued targeted efforts to reduce air pollution in the most affected neighborhoods.
- Health impacts: There is a clear association between PM2.5 exposure and increased hospitalizations for respiratory and cardiovascular diseases, among various age groups. This underlines the significant public health burden associated with fine particulate pollution.

However, the regression analysis performed in this study showed some inconsistencies, resulting in a limitation to this study. This may be due to the limited number of variables considered in the model. To improve the accuracy and predictive power of the analysis, it would be beneficial to include additional variables, such as other respiratory or cardiovascular diseases, socio-economic factors, or environmental conditions that might influence hospitalization rates.

Additionally, future research could benefit from a seasonal analysis of PM2.5 exposure. Since air quality often fluctuates based on seasonal patterns, it would be valuable to investigate whether particular seasons (e.g., winter or summer) are associated with higher PM2.5 levels and, in turn, increased hospitalizations for specific diseases. By breaking down the analysis by season, future studies may uncover more precise relationships between PM2.5 exposure and public health outcomes, allowing for more targeted interventions. Policymakers must continue to prioritize air quality improvements, particularly in areas where high levels of pollution continue to affect public health.

References

Chen, F., Zhang, W., Mfarrej, M. F. B., Saleem, M. H., Khan, K. A., Ma, J., Raposo, A. & Han, H. (2024). Breathing in danger: Understanding the multifaceted impact of air pollution on health impacts. *Ecotoxicology and Environmental Safety*, *280*(116532). https://doi.org/10.1016/j.ecoenv.2024.116532

Christensen, J. (2023, April 19). A quarter of Americans live with polluted air, with people of color and those in Western states disproportionately affected, report says. *CNN Heath*. https://www.cnn.com/2023/04/19/health/state-of-the-air-2023/index.html

Cifuentes, L., Borja-Aburto, V. H., Gouveia, N., Thurston, G., & Davis, D. L. (2001). Assessing the Health Benefits of Urban Air Pollution Reductions Associated with Climate Change Mitigation (2000-2020): Santiago, São Paulo, México City, and New York City. *Environmental Health Perspectives*, *109*, 419–425. https://doi.org/10.2307/3434790

Dockery, D. W. (2009). Health effects of particulate air pollution. Annals of epidemiology, *19*(4), 257–263. https://doi.org/10.1016/j.annepidem.2009.01.018

Goldstein, I. F. (1972). Interaction of Air Pollution and Weather in Their Effects on Health. *HSMHA Health Reports*, *87*(1), 50–55. https://doi.org/10.2307/4594426

Greenburg, L., Jacobs, M. B., Drolette, B. M., Field, F. & Braverman, M. M. (1962). Report of an Air Pollution Incident in New York City, November 1953. *Public Health Reports (1896-1970)*, *77*(1), 7–16. https://doi.org/10.2307/4591399

Gulan, L., Stajic, J. M. & Forkapic, S. (2022). Radon levels and indoor air quality after application of thermal retrofit measures – a case study. *Air Quality, Atmosphere & Health*, *16*, 363-373. https://doi.org/10.1007/s11869-022-01278-w

Kinney, P. L., Chillrud, S. N., Ramstrom, S., Ross, J., & Spengler, J. D. (2002). Exposures to Multiple Air Toxics in New York City. *Environmental Health Perspectives*, *110*, 539–546. http://www.jstor.org/stable/3455248

Lee, J.T., Son, J.Y. & Cho, Y.S. (2007). The adverse effects of fine particle air pollution on respiratory function in the elderly. Science of The Total Environment, *385*(1-3), 28-36. https://doi.org/10.1016/j.scitotenv.2007.07.005

Manisalidis, I., Stavropoulou, E., Stavropoulos & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A Review. *Frontier Public Health, Sec. Environmental Health and Exposome*, *8*. https://doi.org/10.3389/fpubh.2020.00014

Naclerio *et al*. (2020). International expert consensus on the management of allergic rhinitis (AR) aggravated by air pollutants: impact of air pollution on patients with AR: current knowledge and future strategies. *World Allergy Organization Journal*, *13*(100106). https://doi.org/10.1016/j.waojou.2020.100106

Sahu, S. K., Mangaraj, P., Beig, G., Tyagi, B., Tikle, S. & Vinoj, V. (2021). Establishing a link between fine particulate matter (PM2.5) zones and COVID -19 over India based on anthropogenic emission sources and air quality data. *Urban Climate*, *38*(100883). https://doi.org/10.1016/j.uclim.2021.100883

Sicard, P., Agathokleous, E., Anenberg, S. C., Marco, A. D., Paoletti, E., Calatayud, V. (2023). Trends in urban air pollution over the last two decades: A global perspective. *Science of The Total Environment*, *858*(2). https://doi.org/10.1016/j.scitotenv.2022.160064

Simkovich, S.M., Goodman, D., Roa, C., Crocker, M.E., Gianella, G.E., Kirenga, B.J., Wise, R.A., Checkley, W. (2019) The health and social implications of household air pollution and respiratory diseases. *NPJ Primary Care Respiratory Med Medicine*, *29*, 12

Terzano, C., Stefano, F. D., Conti, V., Graziani, E. & Petroianni, A. (2010). Air pollution ultrafine particles: toxicity beyond the lung. European Review for Medical and Pharmacological Sciences, *14*, 809-821. https://www.researchgate.net/profile/Angelo-Petroianni/publication/49743881_Air_pollution_ultrafine_particles_Toxicity_beyond_the_lung/links/02bfe511df1d802388000000/Air-pollution-ultrafine-particles-Toxicity-beyond-the-lung.pdf

United States Environmental Protection Agency: (2024, April 30). *Progress Cleaning the Air and Improving People's Health*. https://www.epa.gov/clean-air-act-overview/progress-cleaning-air-and-improving-peoples-health

US Census Bureau: Population Division. (2020). 2020 Census Detailed Demographic and Housing Characteristics File A (DDHC-A). https://www.nyc.gov/site/planning/planning-level/nyc-population/2020-census.page

Appendix

1. New York City Department of Health, Environment & Health Data Portal. Health impacts of air pollution data. Asthma emergency department visits due to PM2.5. Accessed at https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/health-impacts-of-air-pollution/?id=2117 on 12/01/2024.
2. New York City Department of Health, Environment & Health Data Portal. Health impacts of air pollution data. Respiratory hospitalizations due to PM2.5 (age 20+). Accessed at https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/health-impacts-of-air-pollution/?id=2120 on 12/01/2024.
3. New York City Department of Health, Environment & Health Data Portal. Health impacts of air pollution data. Cardiovascular hospitalizations due to PM2.5 (age 40+). Accessed at https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/health-impacts-of-air-pollution/?id=2117 on 12/01/2024.
4. City of New York. (2020, November 10). Dataset [Air Quality]. NYC OpenData. https://catalog.data.gov/dataset/air-quality
5. Organization for the NYC Environmental Health Services team at DOHMH. (2020). *NYC Geography* [UHF34]. Github. https://github.com/nycehs/NYC_geography
6. Organization for the NYC Environmental Health Services team at DOHMH. (2020). *NYC Geography* [UHF42]. Github. https://github.com/nycehs/NYC_geography
7. Veltman, N. (2018). *Data* [nyc-neighborhoods.geo.json]. Github. https://github.com/veltman/snd3/blob/master/data/nyc-neighborhoods.geo.json?short_path=730a833