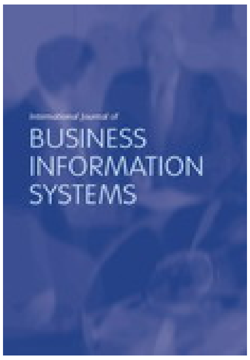


International Journal of Business Information Systems

ISSN online: 1746-0980 - ISSN print: 1746-0972
https://www.inderscience.com/ijbis

Air pollution and asthma hospitalisations in New York City: a borough-level information analysis of spatial, seasonal, and socio-economic determinants

Niranjan Pati, Mayank Anand

DOI: 10.1504/IJBIS.2026.10080580

Article History:

Received:	20 June 2026
Last revised:	26 July 2026
Accepted:	29 July 2026
Published online:	27 August 2026

Air pollution and asthma hospitalisations in New York City: a borough-level information analysis of spatial, seasonal, and socio-economic determinants

Niranjan Pati

Rohrer College of Business,
Rowan University,
NJ, 08028, USA
Email: pati@rowan.edu

Mayank Anand*

College of Science and Mathematics,
Rowan University,
NJ, 08028, USA
Email: mayank.anand3007@gmail.com
*Corresponding author

Abstract: Air pollution drives asthma hospitalisations across New York City's boroughs in ways that are neither uniform nor accidental. This study examines how $PM_{2.5}$, NO_2 and O_3 exposure translates into asthma hospital admissions with borough-level data showing seasonal spikes. The analysis applies quality analytics methods to analyse factors that drive undesirable outcomes. The picture that emerges shows traffic density feeding elevated NO_2, industrial zoning compounding $PM_{2.5}$ loads, and neighbourhood-level poverty reducing the buffer between exposure and severe health consequences rather than pollution causing hospitalisations. Healthcare utilisation dropped sharply in 2020; asthma exacerbation patterns shifted and temporary pollution reductions during lockdowns introduced a confounding variable that the analysis accounts. Indoor exposures, fungal contamination in housing stock and the physical toll of poorly controlled asthma conditions factor in as well. Our findings point toward specific intervention targets where an emission controls near freight corridors and clinic-level early warning systems informed by pollutant monitoring data.

Keywords: particulate; asthma; $PM_{2.5}$; nitrogen dioxide; socio-economic deprivation; quality analytics; Pareto analysis; borough-level variations; seasonal variability.

Reference to this paper should be made as follows: Pati, N. and Anand, M. (2026) 'Air pollution and asthma hospitalisations in New York City: a borough-level information analysis of spatial, seasonal, and socio-economic determinants', *Int. J. Business Information Systems*, Vol. 52, No. 9, pp.1–35.

Biographical notes: Niranjan Pati is a Professor of Management and Entrepreneurship at the Rohrer College of Business at Rowan University, USA. He holds Master's degree and PhD from Northwestern University, Evanston, Illinois. Also, he holds a Master's in Industrial Engineering and Operations Research from Indian Institute of Technology. He served as the Dean of two AACSB accredited business schools, namely, at the Rohrer College of Business, Rowan University (2008–2012) and the School of Business at Indiana University (2001–2008). He published several peer-reviewed journal articles in the areas of operations management, R&D management, technology management, quality management, benchmarking, sustainability, and family business. He has guest-edited special issues of several journals. Currently, he serves as the Chief Editor of the *International Journal of Family Business and Regional Development*, and an associate editor of *Benchmarking: An International Journal*, and serves on the editorial board of three more peer-reviewed journals.

Mayank Anand is a recent graduate of the Master of Science Program in Data Science at Rowan University, Glassboro, New Jersey, USA. He also holds a Master of Computer Applications from Guru Gobind Singh Indraprastha University, New Delhi, and a Bachelor of Computer Applications from Sikkim Manipal University, India. His professional background spans data engineering and applied machine learning, including work at BridgeLabz Solutions, where he developed natural language processing pipelines for survey analytics and delivered technical instruction in Python, distributed computing, and cloud infrastructure, as well as earlier research and analytics roles at The Research World and Analytics Vidhya. His published work spans machine learning for financial risk assessment, including a study on predictive modelling of loan default behaviour for secure banking. His current research interests lie at the intersection of statistical modelling, and applied data science, with particular attention to financial and health systems.

1 Introduction

1.1 Motivation

New York City does not pose a typical urban air quality problem. Population density, freight traffic, industrial operations, and ageing infrastructure compound one another in ways that most cities do not have to manage at the same scale. Long-term exposure to $PM_{2.5}$, NO_2, and O_3 raises the risk of respiratory illness (Espejo et al., 2025; Zhou et al., 2024), and childhood exposure data are especially sobering: prolonged early-life exposure increases the cumulative incidence of asthma in ways that persist well into adulthood (Zanobetti et al., 2024).

Overall pollutant levels in the city have fallen since the late 1990s. Lau et al. (2024) tracked substantial reductions across major pollutants between 1998 and 2021. But borough-level data tells a different story. Areas with heavy freight activities and industrial operations still record pollutant concentrations well above the city average. Pitiranggon et al. (2025) documented sharp $PM_{2.5}$ disparities near highways and freight corridors, and those disparities show up in health outcomes: asthma rates in the Bronx and Brooklyn boroughs consistently exceed citywide figures.

Winter makes things worse. Heating emissions combine with temperature inversions that trap particulates near ground level. Romaszko-Wojtowicz et al. (2025) found that seasonal $PM_{2.5}$ spikes correspond with measurable increases in asthma and respiratory admissions.

This study examines how those spatial and seasonal pollution patterns drive asthma hospitalisation rates across the boroughs and identify where targeted intervention would have the greatest effect. We employ data and visual analytics to apply to environmental health datasets with the aim to recommend measures to improve the quality of life impacted adversely by the incidence of asthma (Alkiayat, 2021). The goal of the study is to converge on findings that are specific enough to act on.

This research is relevant to public health officials, urban planners, environmental policymakers, and healthcare administrators seeking to address air pollution and environmental health disparities in densely populated urban areas. The findings also inform researchers in environmental epidemiology, paediatric pulmonology, and health services research. Community organisations, transportation agencies, and housing authorities may apply these insights to strengthen targeted interventions, improve resource allocation, guide emission reduction strategies, and mitigate neighbourhood level disparities in asthma related health outcomes.

1.2 Identification of the problem

The problem was sourced from two publicly available datasets. Air quality data came from the New York City Community Air Survey (NYCCAS), which provides neighbourhood-level measurements of $PM_{2.5}$, NO_2, and O_3. Asthma hospitalisation records came from the Statewide Planning and Research Cooperative System (SPARCS). Borough-level aggregation across both datasets made it possible to examine seasonal and geographic patterns in pollution exposures and hospital admissions adjacent to each other.

The pattern that emerged from data cleaning and integration was clear: boroughs with higher pollutant concentrations also had higher asthma hospitalisation rates. The Bronx and Brooklyn boroughs were the most consistent examples. Traffic density, industrial emissions, and ageing infrastructure all push air quality below the city average there. Shukla et al. (2022) documented significant variation in particulate matter exposures at the ZIP-code level across boroughs, which supports the case for granular geographic analysis rather than city-wide aggregation when the question is posed as to who bears the health burden.

Temporal analysis of seasons shows recurring winter $PM_{2.5}$ spikes resulting in increases in asthma admissions. Winter months intensify pollutant accumulation due to low temperatures that create increased heating emissions. Hospital admission data reflects that, consistent with Romaszko-Wojtowicz et al. (2025).

1.3 Characterisation of the problem

Air pollution in New York City is not one problem, but several layered-on top of each other. Environmental exposure, healthcare outcomes, and structural inequality are all involved, and they are hard to separate. $PM_{2.5}$, NO_2, and O_3 remain elevated in neighbourhoods with heavy traffic, industrial activities, and limited green space. The damage extends beyond asthma. Increased mortality, cardiovascular disease, and respiratory illness have all been linked to pollution exposure in the city (Goldberg et al., 2025; Humphrey et al., 2024; Mananga et al., 2023).

Winter makes things predictable but still in deleterious ways. Heating emissions rise, wind circulation drops, and atmospheric stagnation traps pollutants near ground level. These conditions can push a stable asthma patient into an emergency admission within a day or two.

This study applies data and visual analytics tools to derive evidence base via trend analysis, Pareto charts, correlation analysis and root cause identification. The approach draws on multi-criteria assessment methods increasingly used in healthcare sustainability research (Thomas and Suresh, 2025). The goal is to identify which pollution sources and borough-level conditions account for the largest share of asthma hospitalisations and where reducing that burden can be achieved.

2 Literature review

The link between ambient air pollution and asthma hospitalisation in New York City is well documented (Zhou et al., 2024; Espejo et al., 2025). $PM_{2.5}$ and NO_2, concentrated along freight corridors and in neighbourhoods with ageing infrastructure, are consistently associated with elevated asthma emergency department visits and admissions. The ZIP code-level analysis confirms what borough level aggregation suggests: particulate matter exposures and asthma morbidity cluster occur in the same low-income and minority communities (Shukla et al., 2022).

Winter is the worse season for pollution. Heating emissions and atmospheric stagnation push $PM_{2.5}$ higher, and increased hospitalisation rates follow. This alarming trend is conspicuous, particularly among children and the elderly (Romaszko-Wojtowicz et al., 2025). Indoor conditions add to the burden. Cochran et al. (2022) found elevated allergenic fungal concentrations in household dust during spring months in paediatric cohorts across the city, suggesting that seasonal exacerbation has both outdoor and indoor dimensions. The data shows long-term average pollution levels have improved, but short-term spikes still produce acute hospitalisation increases. Thus, acceptable averages do not eliminate episodic risks (Mananga et al., 2023).

The June 2023 Canadian wildfire smoke event brought an undesirable twist. Emergency department visits for asthma rose measurably across the city during that episode (Chen et al., 2023; Thurston et al., 2023). Transboundary pollution events can generate acute respiratory burdens within days, and local monitoring systems are not equipped to catch them in time. Real-time surveillance capable of tracking episodic exposures is an operational gap. Failures of infrastructure pose similar risks. Flores et al. (2026) found that electrical power outages in New York City were associated with measurable increases in asthma-related emergency department visits between 2018 and 2022, suggesting that episodic environmental disruptions extend beyond air quality events alone.

Attributing outcomes to individual pollutants is harder than it looks. $PM_{2.5}$, NO_2, and O_3 tend to co-vary spatially and temporally, which makes isolating any single pollutant's contribution methodologically difficult. Humphrey et al. (2024) encountered the same problem working on cardiovascular outcomes. Thus, multivariate models that account for co-exposure are necessary to put the finger on the pulse of inferring reliably about which pollutants are doing the most damage (Zhou et al., 2024; Espejo et al., 2025).

Socio-economic factors cut across all of this. Poverty, minority status, and insurance coverage are often stronger predictors of asthma hospitalisation in statistical models than pollutant concentrations alone (Shukla et al., 2022; Mananga et al., 2023). Higher pollution exposure and reduced healthcare access tend to occur in the same neighbourhoods, and each adversely impacts the effect of the other.

The long-term air quality picture in New York City is broadly positive. Monitoring data from 1998 to 2024 shows sustained reductions in $PM_{2.5}$ and NO_2 following regulatory interventions. These measurable improvements in population-level respiratory and cardiovascular health provide supporting evidence on the improved mortality level (Johnson et al., 2020; Lau et al., 2024; Goldberg et al., 2025). The gains are real, but they have not been evenly distributed across the boroughs. Low-income and minority neighbourhoods still carry higher pollution burdens. Episodic events such as wildfire smoke or winter stagnation remain a threat regardless of where the long-term averages land.

Literature pertaining to the study has three consistent gaps. Most studies track short-term emergency visits rather than multi-year hospitalisation trends. Many rely on citywide pollution averages that mask what is happening at the borough level. And few combine environmental and socio-economic data in ways that could show whether air quality improvements have reached the communities most affected.

This study explored the borough level environmental and demographic data over a long-time horizon. The central question is not just whether pollution and asthma go in tandem but whether the boroughs that have historically borne the greatest burden have seen their burden ease.

3 Research methodology

This study applied data and visual analytics tools to analyse air pollution and asthma hospitalisation with the objective to improve the quality of life across New York City's boroughs. The core questions addressed by the study were

1 Where and when are pollution levels the highest?

2 How do they relate to hospitalisation rates?

3 Which structural and demographic factors moderate that relationship?

Tool selection followed from those questions and from the quantitative, time-series, and spatial nature of the available data. Data visualisation methods of this kind have been shown to strengthen decision-making in other applied domains as well (Ghozlan and Shanab, 2025).

$PM_{2.5}$ and NO_2 are the primary independent (exposure) variables, chosen for their established links to respiratory morbidity in urban settings and their relevance to both traffic-related and stationary emission sources across the boroughs. Statistical analysis establishes quantitative relationships between pollutant concentrations and hospitalisation rates. Affinity diagrams and interrelationship digraphs support qualitative assessment and causal mapping. Methods were piloted on a data subset before full application to confirm that they provided the fit for the purpose. Borough-level indicators of poverty, housing density, and land use are included as covariates alongside environmental exposure data.

The study period spans pre-pandemic and COVID-19 years. The pandemic disrupted healthcare utilisation altered pollution profiles, and shifted asthma exacerbation trends in ways that complicated time-series interpretation. Pandemic era observations are examined for structural breaks, consistent with the behavioural and environmental changes of early 2020, and findings from that period are read with that disruption in view.

Control charts track variations in pollution concentrations and hospitalisation rates over time, flagging statistically significant deviations and seasonal patterns. Pareto analysis identifies which sources, and geographic areas account for the largest share of the asthma burden. On the straightforward logic those intervention resources should go where the evidence says harm is concentrated (Deepu and Ravi, 2026). Interrelationship digraphs map causal dependencies among pollution sources, socio-economic conditions, indoor environments, and health outcomes. Together they provide a system-level account of what drives asthma hospitalisations across the city, one that statistical modelling alone would not unravel.

4 Data analysis

4.1 Sources

Air quality data came from the NYCCAS. Borough-level $PM_{2.5}$ measurements were retrieved as CSV files covering 2019 to 2026. A broader pollutant series, $PM_{2.5}$, NO_2, and O_3, was drawn from yearly charts spanning 2009 to 2025; 2026 values were not directly available in that series and were estimated using linear regression fitted to the 2009–2025 trend data. All files were converted into the standardised CSV format. The dataset includes winter, summer, and annual summaries, and the combination of the borough-level records and the regression-extended chart series covers enough temporal range to examine both long-term air quality shifts and seasonal pollution patterns.

Asthma hospitalisations data came from the SPARCS, which holds comprehensive hospital discharge records for New York State. The SPARCS extract covers 2010 to 2024 and includes asthma cases, emergency department visits, and hospitalisation costs, aggregated by borough and year. These records were cross validated against the NYC Environment and Health Data Portal.

Demographic and socio-economic data came from the American Community Survey (ACS), with annual borough-level indicators from 2010 to 2024. The three datasets do not share identical temporal coverage, so integrated analyses linking pollution exposure, asthma outcomes, and demographic variables are restricted to 2010–2024, the period where all three overlap. Full variable definitions, spatial resolution, temporal coverage, and source file formats are listed in Appendix, Table A1.

4.2 Exploratory analysis

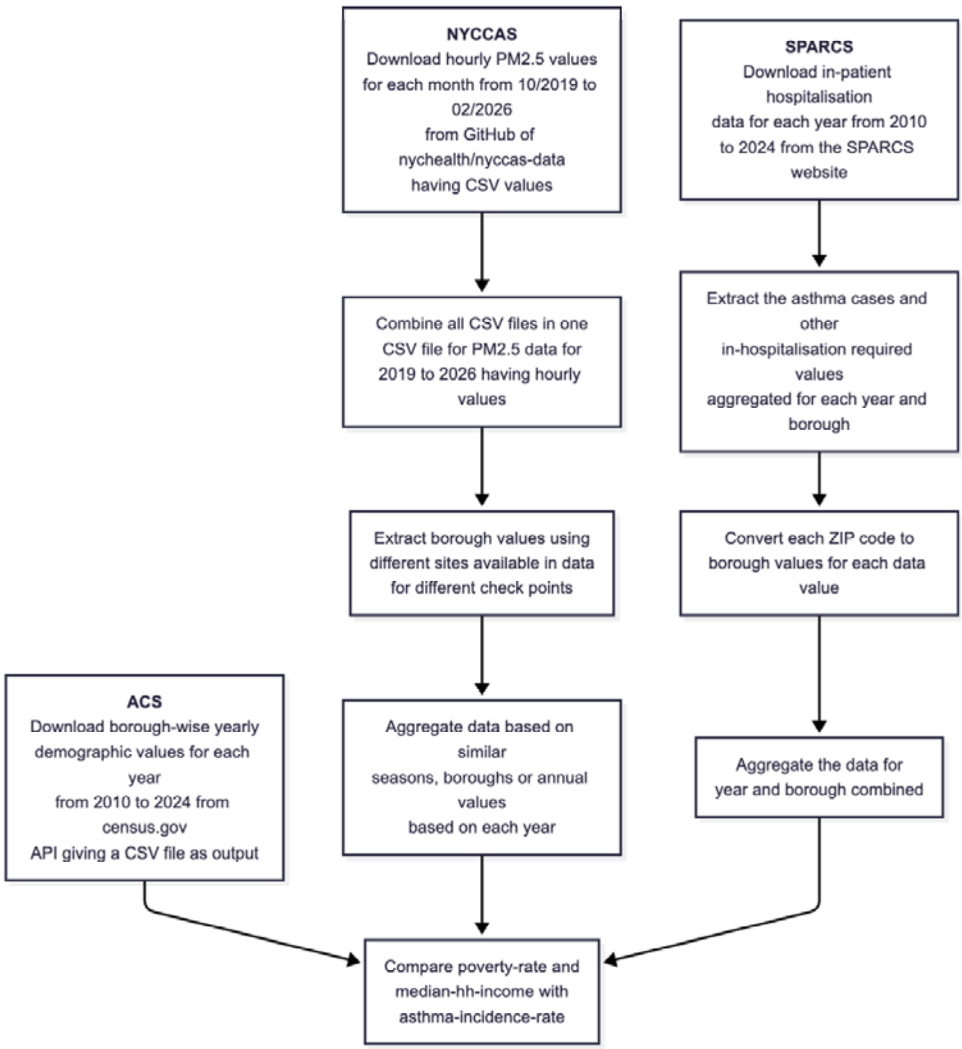
4.2.1 Data collection process

All source datasets were downloaded in their original formats and stored in a version-controlled repository. Metadata logs recorded the source, acquisition date, file names, and original record structures for each file. Raw data were not modified at any point, so every transformation step is traceable, and the full pipeline can be reconstructed from source through to final analytical datasets.

Three datasets fed the exploratory analysis: NYCCAS air quality monitoring data, SPARCS hospital discharge records, and ACS socio-economic indicators.

Historic NYCCAS CSV files were programmatically downloaded, combined, and aggregated (Appendix, Notebook S1). SPARCS extracts were filtered to isolate asthma-related diagnoses and borough-level hospitalisation counts (Appendix, Notebook S3). ACS demographic variables were retrieved annually for 2010–2024 (Appendix, Notebook S2). Intermediate outputs were saved at each stage of the workflow (Appendix, Table A3).

Figure 1 Data collection and integration workflow for the NYC asthma analysis



Source: Our own (Notebooks S1–S7; Table A2)

4.2.2 Initial data merging

Datasets were merged before extensive cleaning or filtering. The rationale was to avoid discarding information prematurely. Cleaning decisions made on isolated datasets can remove records that become meaningful once sources are joined. Merging first also supported cross-source validation and made subsequent imputation and anomaly detection more reliable.

The process followed four stages. Join keys were standardised first: borough names converted to title case, year fields converted to integers, seasonal records aligned to standard quarterly definitions. NYCCAS monitoring sites were mapped to boroughs to produce borough-level pollutant averages that could be joined to SPARCS and ACS data.

Pollutant naming conventions, temporal fields, and demographic identifiers were then harmonised across all three datasets. Hierarchical merges followed in the following sequence. NYCCAS seasonal and annual summaries first, SPARCS hospitalisation aggregates second, ACS demographic indicators third. Outer joins were used throughout to retain incomplete observations rather than drop them at the merge stage.

Each merged row retained references to its source files and notebook outputs, keeping the pipeline traceable at the record level. The resulting master dataset underpinned all subsequent exploratory analyses, statistical summaries, and visual diagnostics.

4.2.3 Data pre-processing and cleaning

All cleaning was applied to the merged dataset after integration, covering missingness, imputation, standardisation, duplicate handling, and outlier detection.

Missingness was assessed at both the variable and dataset level. Variables with systematic missingness above 30% were flagged for conditional exclusion. ACS socio-economic variables were carried forward or backward from the nearest available year where appropriate (Appendix, Notebook S2).

Imputation varied by variable types. Seasonal pollutant gaps were addressed using a three-step hierarchy. The borough-season long-run means first, borough-year interpolation for contiguous series second, and nearest-borough sensor borrowing where values remained missing and were needed downstream. SPARCS hospitalisation counts were treated more conservatively. Borough-year observations with missing asthma counts were kept as missing and excluded from incidence-rate denominators.

Borough names, pollutant naming conventions, seasonal indicators, and temporal fields were standardised across the merged dataset. Duplicate records on complete join keys were collapsed by averaging numeric fields or reconciled using quality indicators where available. All duplicate resolution actions were logged.

Outlier detection applied IQR thresholds and Z-score analysis to log-transformed pollutant variables. Suspected anomalies were cross validated against source indicators and contextual data to distinguish real environmental events from sensors or recording errors. Outliers were flagged with an outlier flag variable rather than removed. This kept them available for sensitivity analyses without letting them distort primary results.

4.2.4 Exploratory summaries and diagnostics

Exploratory analysis combined statistical summaries, control charts, scatter plots, regression diagnostics, and short-term forecasting to examine temporal, geographic, and socio-economic patterns across the dataset.

Seasonal $PM_{2.5}$ patterns from NYCCAS data were summarised using control charts. The grand mean across the combined seasonal sequence was approximately 9.00 μg/m^3, with upper and lower control limits of 15.87 μg/m^3 and 2.13 μg/m^3. Seasonal line plots identified recurring fluctuations and periods approaching those thresholds.

NO_2 and O_3 were analysed the same way. NO_2 had an overall seasonal mean of 21.18 ppb, with control limits of 35.55 ppb and 6.81 ppb. Summer O_3 averaged 30.47 ppb, with limits of 36.49 ppb and 24.44 ppb. Short-term forecasting tested four models. They were linear regression, polynomial regression (degree 2), decision tree regression, and exponential smoothing, evaluated by RMSE over the final three years using train-test splits. Polynomial transformations were included to improve flexibility on small datasets.

Borough-level control charts showed that several boroughs intermittently exceeded control limits, pointing to different pollution dynamics across the city rather than a uniform citywide pattern.

Asthma incidence rates per 100,000 residents were calculated as:

$$Asthma\ incidence\ rate = (Hospital\ asthma\ cases / ACS\ \ population) \times 100{,}000$$

The Bronx stood apart, due to its incidence rates frequently exceeding 250–300 cases per 100,000 in later years in the data set, well above every other borough. Manhattan and Staten Island were consistently at the lower end, typically between 60–110 cases per 100,000. Results were saved as merged_asthma_data.csv.

Scatter plots and regression diagnostics examined pollutant-health and pollutant-socio-economic relationships at the borough level (Appendix, Notebooks S4–S5). $PM_{2.5}$ and asthma incidence showed positive spatial clustering. The Bronx, combining high pollution with high asthma rates, was the glaring example. Regression on ACS socio-economic variables found a negative correlation between median household income and asthma incidence, and a positive one between poverty indicators and asthma burden. Correlation coefficients were computed across both workflows to quantify the direction and relative strength of each relationship.

4.2.5 Feature engineering

Following the exploratory analysis, several variables were engineered for use in statistical modelling, forecasting, and quality diagnostics. Derived features included borough × year seasonal indicators, binary seasonal flags, and a composite pollution index combining $PM_{2.5}$, NO_2, and O_3 concentrations weighted by health impact.

Quality-control and provenance flags were added alongside these: imputation status, outlier identification, and source-tracking variables recording the origin and quality status of each observation (Appendix, Table A2). Borough-level annual and seasonal pollutant averages were derived to feed control-chart generation, regression analysis, and borough-level comparisons in subsequent stages.

All principal engineered datasets are documented in Appendix, Table A3.

4.2.6 Final integration, quality control, and export

After cleaning and feature engineering, a final quality-control pass checked dataset integrity and analytical consistency. This covered impossible or negative pollutant concentrations, alignment between derived asthma incidence rates and raw hospitalisation counts, and recalculation of borough-level summary statistics and control limits.

The exploratory pipeline exported several datasets used in subsequent analyses and reporting. They have been documented in Appendix, Table A3. Each notebook records its derivation steps and can be re-executed to reproduce all outputs (Appendix, Notebooks S1–S5).

The reporting workflow produced control-chart visualisations, regression and scatter plots, and forecasting comparison tables with RMSE performance across modelling approaches for NO_2 and O_3. These outputs fed into identification of seasonal pollution patterns, borough-level outliers, and pollutant-health relationships examined in the formal modelling stage.

Merge and cleaning logs were retained alongside analytical outputs. Per-file provenance records document source datasets, acquisition dates, transformation stages, and record counts for each file in the pipeline.

4.3 Cost-benefit considerations

This section examines the economic trade-offs from the perspective of public decision-makers relative to what does it cost to reduce air pollution in New York City's high-burden boroughs, and how does the inaction impact the city in terms of health and cost?

4.3.1 Costs of pollution mitigation strategies

The main costs fall into several categories. Monitoring infrastructure means installing and maintaining additional sensors in pollution hotspots. Regulatory enforcement translates into expanded inspection of vehicle and industrial emissions. Transition incentives cover subsidies for cleaner heating fuels and low-emission transport. Traffic management measures such as recently introduced congestion pricing, low-emission zones, etc. carry their own implementation costs, as do green infrastructure investments such as roadway buffer zones and urban vegetation.

These interventions are geographically targeted, which matters. Resources can be concentrated in the boroughs where pollution burdens and health risks are greatest rather than spread thinly across the city.

4.3.2 Benefits of pollution reduction

The most direct benefit is fewer asthma hospitalisations and emergency department visits that result in lower healthcare costs and reduced pressure on the hospital system. Even modest reductions in $PM_{2.5}$ and NO_2 in high-exposure boroughs are likely to reduce asthma exacerbations, particularly among children and the elderly. Indirect gains include improved productivity, lower school and work absenteeism, and reduced long-term chronic disease burden.

There is also a distributional argument. The boroughs that bear the heaviest pollution burden are not the ones that generated it. Reducing that disparity is not only an economic calculation but also a social imperative.

4.3.3 Net value for decision-makers

Targeted intervention in a small number of high-burden boroughs can deliver large public health gains relative to cost. The analytical framework developed here identifies where investment is likely to produce the greatest reduction in asthma hospitalisations, affording decision-makers the basis for allocation rather than requiring them to act on intuition or political pressure alone. It is not gainsaying the fact that data-driven pollution control delivers better outcomes than uniform citywide approach, which spread resources more evenly but produces smaller marginal gains where the need is greatest.

5 Quality analytics

Four quality analytics tools were applied to the NYC air quality and asthma hospitalisation dataset: descriptive statistics and trend analysis, affinity diagrams, interrelationship digraphs, and control charts. Tool selection followed from the quantitative, seasonal, and geographically segmented structure of the data and from the research objective of examining relationships between environmental pollution and public health outcomes. The tools applied are summarised in Table A4.

5.1 Descriptive statistics and trend analysis

Descriptive statistics captured by mean, median, standard deviation, and range, established baseline pollution conditions across boroughs and seasons, and identified where $PM_{2.5}$, NO_2, and O_3 concentrations were consistently elevated. Trend analysis using line graphs and seasonal comparisons showed recurring fluctuations in pollution concentrations, with winter periods consistently producing highest readings.

5.2 Affinity diagram

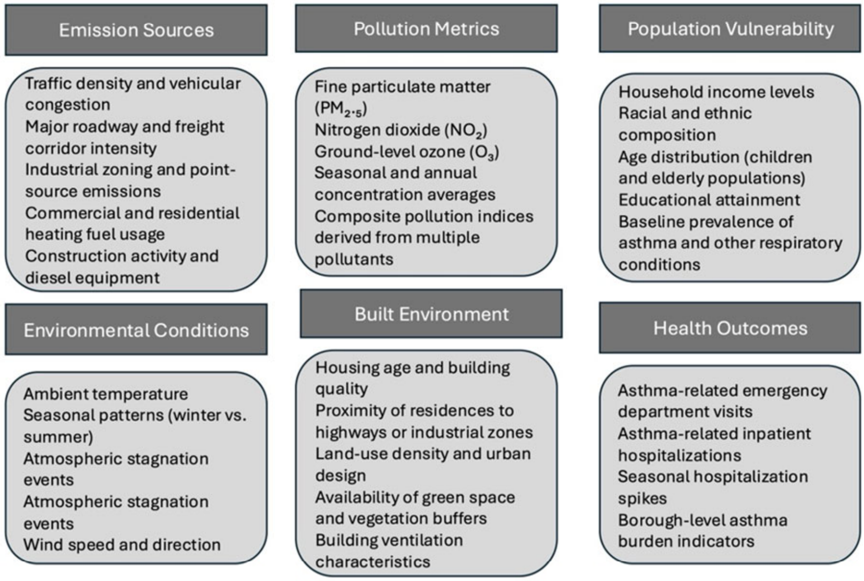
An affinity diagram organised the dataset variables into three clusters: pollution sources, environmental metrics, and health outcomes. This grouping clarified the relationships among emissions, ambient air quality indicators, and asthma hospitalisations, and laid the groundwork for the causal analysis that followed.

5.3 Interrelationship digraph

The interrelationship digraph mapped how traffic and industrial emissions drive ambient pollution levels, which in turn influence asthma hospitalisation rates across boroughs. Emissions control emerged from this analysis as the highest-leverage intervention point, the node with the most outgoing causal arrows and the greatest downstream effect on health outcomes.

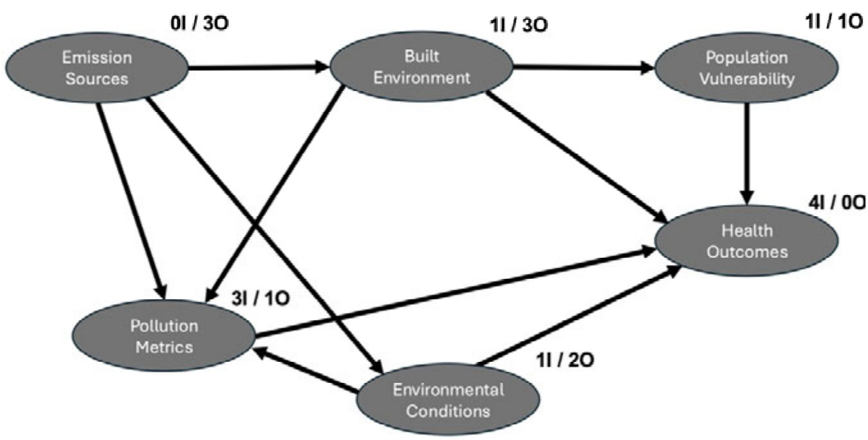
The causal mapping also generated three hypotheses. Boroughs with consistently high $PM_{2.5}$ and NO_2 concentrations are expected to show significantly higher asthma hospitalisation rates after adjusting for population size. Seasonal $PM_{2.5}$ increases, particularly in winter, are expected to correspond with seasonal increases in asthma admissions. And traffic density and industrial zoning likely act as indirect drivers of hospitalisation through their contribution to local NO_2 levels.

Figure 2 Affinity diagram categorising variables into pollution sources, environmental metrics, and health outcomes



Source: Our own

Figure 3 Interrelationship digraph of causal dependencies between traffic emissions, industrial activity, ambient pollution levels, and asthma hospitalisation rates across New York City boroughs



Source: Our own

5.4 Control charts

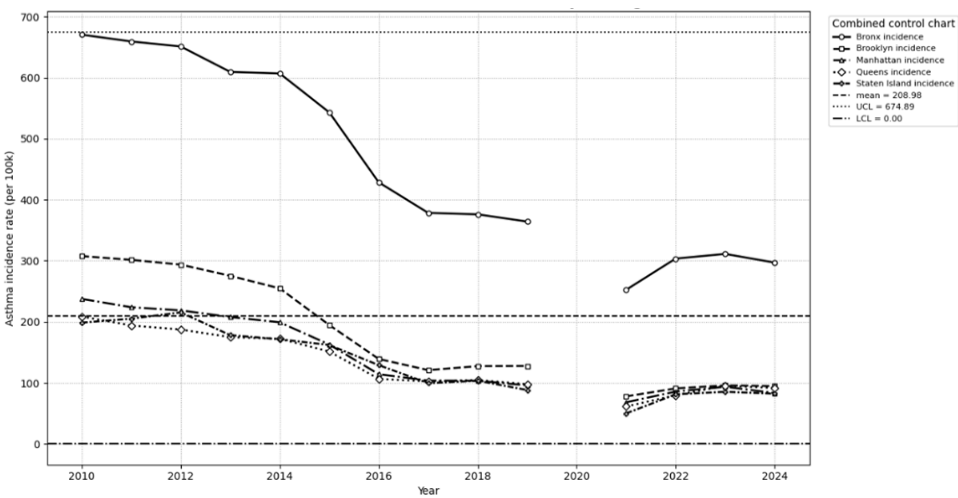
Control charts assessed process stability in both asthma incidence rates and pollution concentrations and flagged significant deviations for further investigation.

Figure 4 tracks asthma hospitalisation rates per 100,000 residents across all five boroughs from 2010 to 2024, with a grand mean of 208.98, UCL of 674.89, and LCL of

0.00 (calculated using $\bar{X} \pm 3\sigma$). There is no data point for 2020 because hospitals were not reliably reporting non-COVID admissions during the pandemic. So, the gap reflects a measurement problem rather than any genuine improvement in outcomes.

The Bronx dominates the chart throughout. Rates were above 600 per 100,000 in the early 2010s and fell to around 300 by 2024. This can be construed as a real progress for Bronx, but still roughly three times the process means and well above every other borough. All five boroughs showed a downward trend over the period. The Bronx has not closed the gap, which points to the fact that system-wide improvement has escaped this borough.

Figure 4 Combined asthma incidence control chart by borough

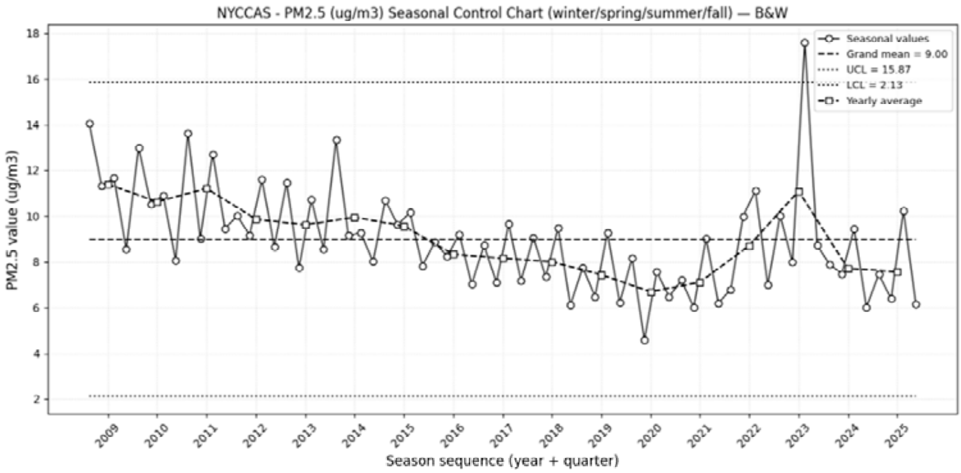


Source: Our own (Appendix, Notebook S3; input: SPARCS-data.csv)

The $PM_{2.5}$ seasonal control chart (Figure 5) plots all quarterly measurements as a continuous sequence from 2009 to 2025, with a grand mean of 9.00 μg/m^3, UCL of 15.87 μg/m^3, and LCL of 2.13 μg/m^3. The overlaid yearly average trend line runs from approximately 10.5 μg/m^3 in 2009 to 7.5 μg/m^3 by 2025, a sustained downward trajectory reflecting cumulative emissions reductions. Two observations stand out. A summer 2023 spike reached approximately 17.5 μg/m^3, breaching the UCL and consistent with the Canadian wildfire smoke episode. A 2020 reading approached the LCL at approximately 4.6 μg/m^3, consistent with the sharp drop in traffic and industrial activity during the pandemic. All other seasonal values stayed within control limits.

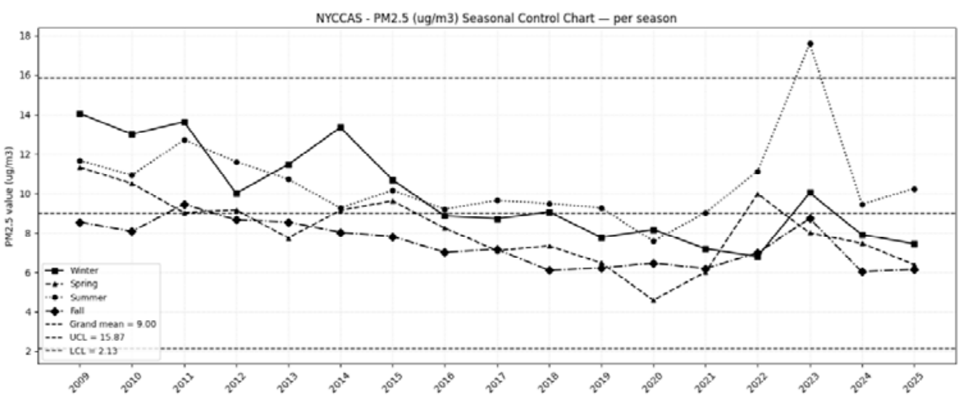
The per-season chart (Figure 6) separates $PM_{2.5}$ trajectories by season across the same period. Winter consistently recorded the highest concentrations, peaking at approximately 14 μg/m^3 in 2009 and 13.5 μg/m^3 in 2011 and 2014, before declining to around 7.5–8 μg/m^3 by 2024–2025. Summer 2023 produced the only UCL breach across any season at approximately 17.5 μg/m^3. This is due to the Canadian wildfire episode. Spring and fall followed similar downward paths, with fall values declining from approximately 8.5 μg/m^3 in 2009 to around 6 μg/m^3 by 2025. All four seasons converged below the grand mean of 9.00 μg/m^3 in later years.

Figure 5 NYCCAS $PM_{2.5}$ ($\mu g/m^3$) Seasonal control chart displaying quarterly winter, spring, summer, and fall values with yearly average trend, 2009–2025



Source: Our own (Appendix, Notebook S4; input: NYCCAS-data.csv)

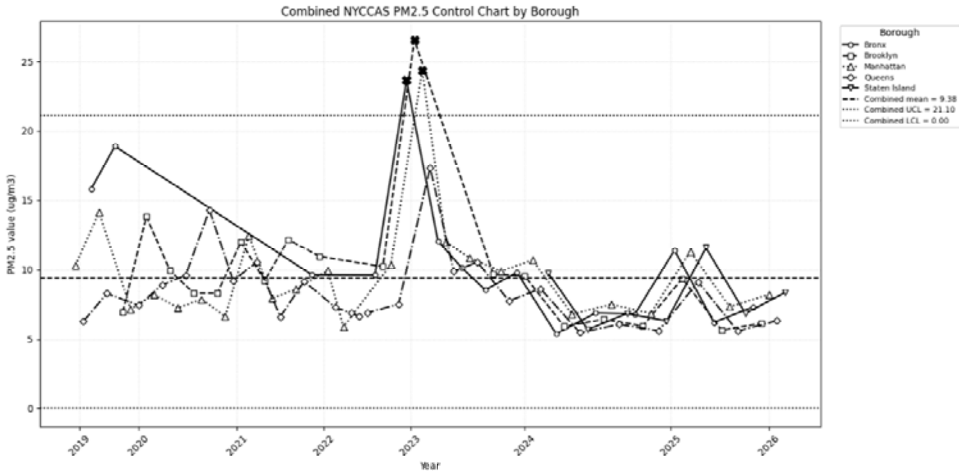
Figure 6 NYCCAS $PM_{2.5}$ ($\mu g/m^3$) per-season control chart displaying separate winter, spring, summer, and fall trajectories, 2009–2025



Source: Our own (Appendix, Notebook S4; input: NYCCAS-data.csv)

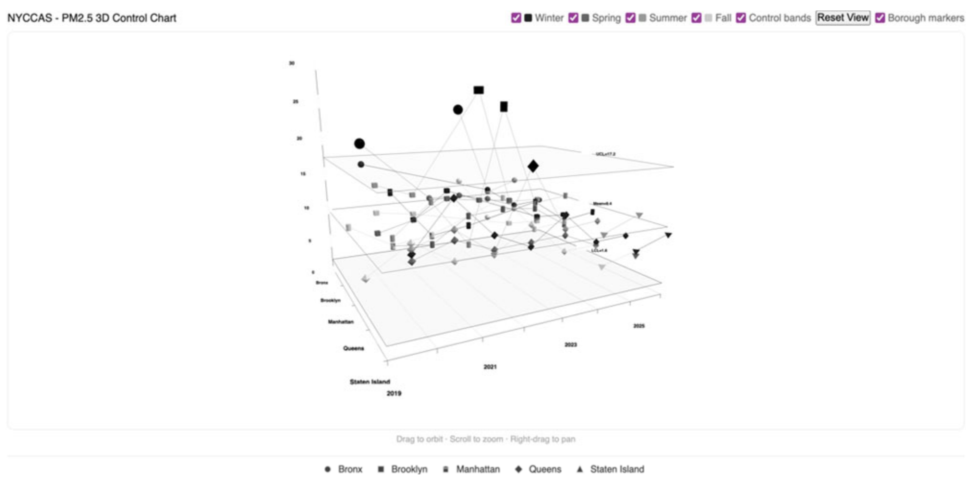
Figure 7 disaggregates $PM_{2.5}$ by both borough and season across 2019–2026, plotting four seasonal readings per borough per year. The Bronx records the highest seasonal values across all four seasons throughout the period. Its winter readings reflect heating emissions and atmospheric stagnation; its summer 2023 spike, alongside Brooklyn's, breached the combined UCL of 21.10 $\mu g/m^3$, the only out-of-control observations in the dataset, and both attributable to the Canadian wildfire episode. Brooklyn's seasonal pattern closely mirrors the Bronx at slightly lower absolute concentrations. Queens and Manhattan show moderate seasonal variation, with winter readings periodically approaching the UCL in earlier years before declining toward the mean by 2024–2026. Staten Island records the lowest values across all boroughs and seasons, staying well within control limits throughout.

Figure 7 NYCCAS $PM_{2.5}$ ($\mu g/m^3$) borough-specific control chart displaying separate Bronx, Brooklyn, Manhattan, Queens and Staten Island trajectories per borough, 2019–2026



Source: Our own (Appendix, Notebook S4; input: NYCCAS-Seasonal-Borough.csv)

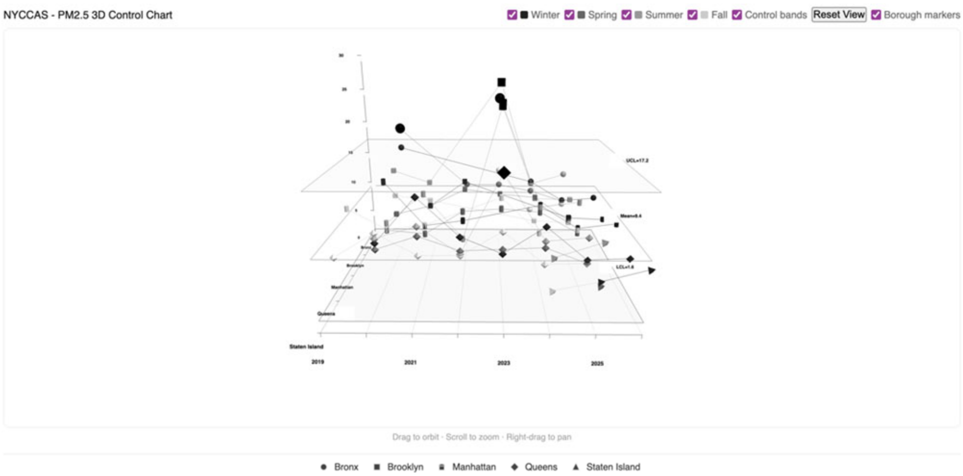
Figure 8 NYCCAS 3D borough-seasonal $PM_{2.5}$ control chart: front-elevation view showing borough axis, year axis, and UCL/mean/LCL control planes (see online version for colours)



Source: Our own (Appendix, Notebook S4; input: NYCCAS-Seasonal-Borough.csv)

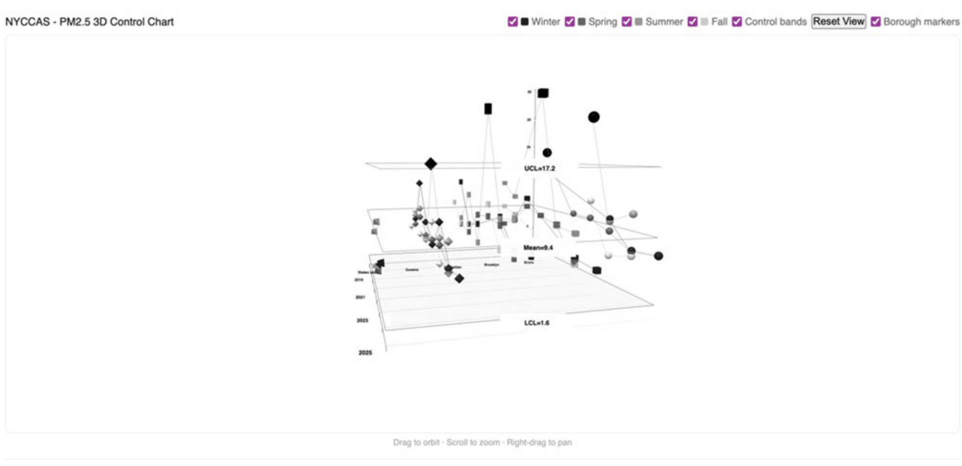
Several cross-borough patterns are clearer in the seasonal disaggregation than in annual averages. The summer 2023 anomaly is the sharpest: Bronx and Brooklyn summer readings exceed the UCL while the same season for Queens, Manhattan, and Staten Island (though elevated relative to their own prior readings) stay below the threshold. The wild-fire episode imposed a disproportionate additional burden on boroughs that were already above average. A secondary pattern is the compression of seasonal variation in later years: the wide within-borough swings between winter highs and summer lows visible in 2019–2021 gave way to lower, more stable readings across all four seasons by 2024–2026. The pandemic year of 2020 shows atypically low spring and summer readings across most boroughs, consistent with suppressed traffic and commercial activity, before values rebounded in 2021.

Figure 9 NYCCAS 3D borough-seasonal $PM_{2.5}$ control chart: oblique view highlighting out-of-control points (red spheres) relative to the UCL plane, 2019–2026 (see online version for colours)



Source: Our own (Appendix, Notebook S4; input: NYCCAS-Seasonal-Borough.csv)

Figure 10 NYCCAS 3D borough-seasonal $PM_{2.5}$ control chart, side-elevation view showing temporal trends across boroughs from 2019 to 2026 (see online version for colours)



Notes: Static views are presented above. The fully interactive version, built using visual analytics principles to support 3D rotation, zoom, and seasonal filtering, is available at https://doi.org/10.5281/zenodo.20221725.

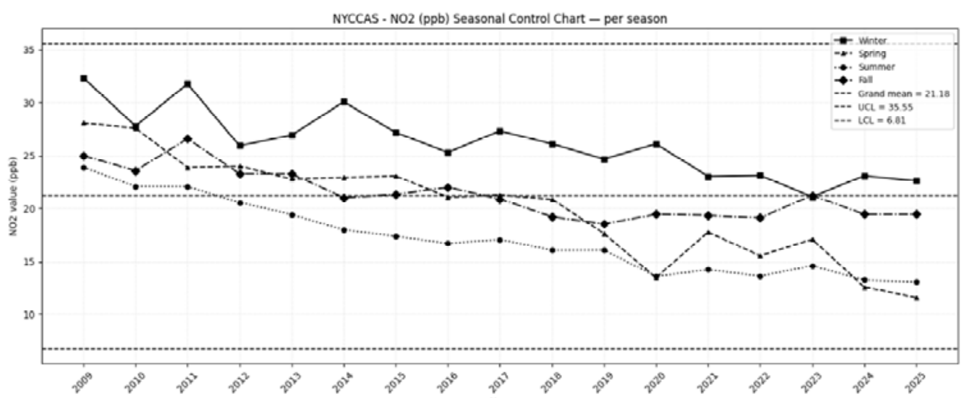
Source: Our own (Appendix, Notebook S4; input: NYCCAS-Seasonal-Borough.csv)

An interactive 3D control chart was also developed to examine $PM_{2.5}$ variation across boroughs, seasons, and years simultaneously (Figures 8, 9 and 10). Each borough-season-year observation is plotted as a sphere in three-dimensional space, with control planes at the UCL (17.2 μg/m^3), mean (9.4 μg/m^3), and LCL (1.6 μg/m^3). Seasonal observations are colour-coded (blue = winter, green = spring, orange = summer, red = fall), and out-of-control points are highlighted in red.

The three perspectives show the same story from different angles. Bronx and Brooklyn observations cluster above the mean plane and breach the UCL during the summer 2023 wildfire episode and in earlier high-pollution winters. Queens, Manhattan, and Staten Island sit predominantly below the mean plane. The Bronx's status as a spatial outlier is immediately apparent in the 3D rendering. Its observations occupy a different region of the chart from every other borough.

The NO_2 seasonal control chart (Figure 11) covers 2009 to 2025, with a grand mean of 21.18 ppb, UCL of 35.55 ppb, and LCL of 6.81 ppb. Winter consistently recorded the highest NO_2 levels across all years. Summer values showed the sharpest decline, falling from approximately 24 ppb in 2009 to around 13 ppb by 2025. All seasonal values remained within control limits throughout indicating a stable process with a consistent downward trend. The persistent winter elevation reflects heating emissions and reduced atmospheric dispersion in colder months.

Figure 11 NO_2 seasonal control chart

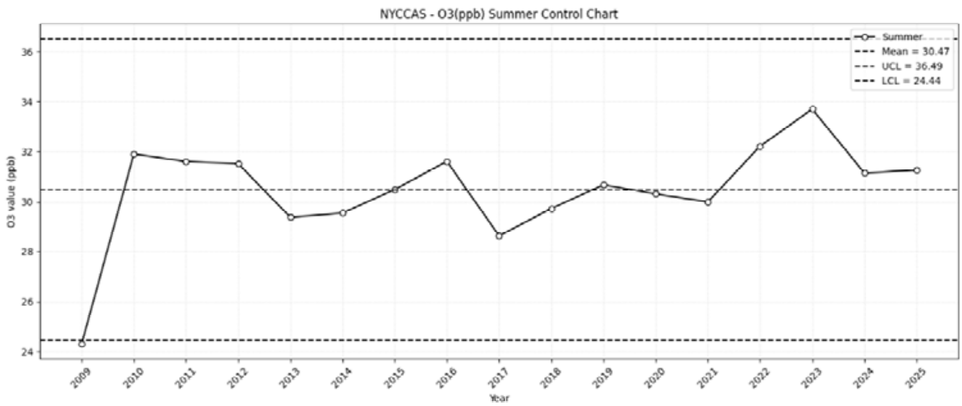


Source: Our own (Appendix, Notebook S4; input: NYCCAS-data.csv)

The O_3 summer control chart (Figure 12) covers summer months only, with a mean of 30.47 ppb, UCL of 36.49 ppb, and LCL of 24.44 ppb. Unlike $PM_{2.5}$ and NO_2, ozone shows no consistent downward trend. A low of approximately 24.3 ppb in 2009 approached the LCL; a peak of 33.8 ppb in 2023 approached the UCL. The absence of improvement is not surprising. This is because ozone forms through photochemical reactions involving sunlight and precursor gases rather than through direct emissions, and the interventions that have reduced $PM_{2.5}$ and NO_2 do not impact it in the same way.

Across the control charts, air quality in New York City has improved on most pollutants and across most boroughs over the study period. But certain geographies and seasons continue to produce elevated exposures, and episodic events such as wildfire smoke, winter stagnation, can push readings out of control regardless of the long-term trend. Continuous monitoring and integration of supplementary contextual data, including weather records and NYC DEP alerts, remain necessary for timely root-cause identification.

Figure 12 O_3 in summer control chart



Source: Our own (Appendix, Notebook S4; input: NYCCAS-data.csv)

5.5 Pareto analysis of asthma hospitalisation

The Pareto chart (Figure 13) ranks four hospitalisation reason groups by total count. Respiratory infection leads at around 145,000 hospitalisations, followed by respiratory failure (approximately 86,000), other respiratory disorders (roughly 81,000), and asthma exacerbation (approximately 58,000). The cumulative line crosses the 80% threshold between the third and fourth bars: the first three categories together account for around 82% of all admissions.

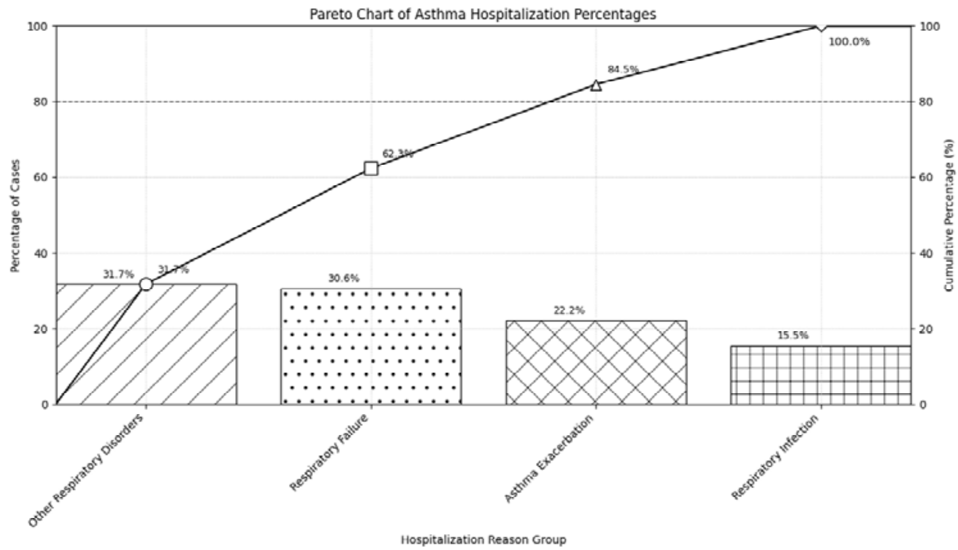
Asthma exacerbation, despite being the most direct clinical trigger, ranks last. That ranking has practical implications for intervention planning. Targeting exacerbation events alone addresses under a fifth of the total admission burden. Respiratory infections drive far more hospital visits in asthma patients. For the subset of patients who deteriorate to respiratory failure, clinical management choices such as non-invasive ventilation have shown to meaningfully affect outcomes (Buss et al., 2026), underscoring that upstream prevention and downstream clinical care both matters. This calls for infection prevention by instituting vaccination uptake, early antibiotic stewardship, etc. as a higher-volume target than optimising bronchodilator or corticosteroid pathways alone. These priorities are discussed in Subsection 5.6.

5.6 Process decision program chart

A process decision program chart (PDPC) mapped implementation risks for four proposed interventions: urban green infrastructure expansion, boiler modernisation, congestion pricing and low-emission traffic controls, and an asthma alert and public education system.

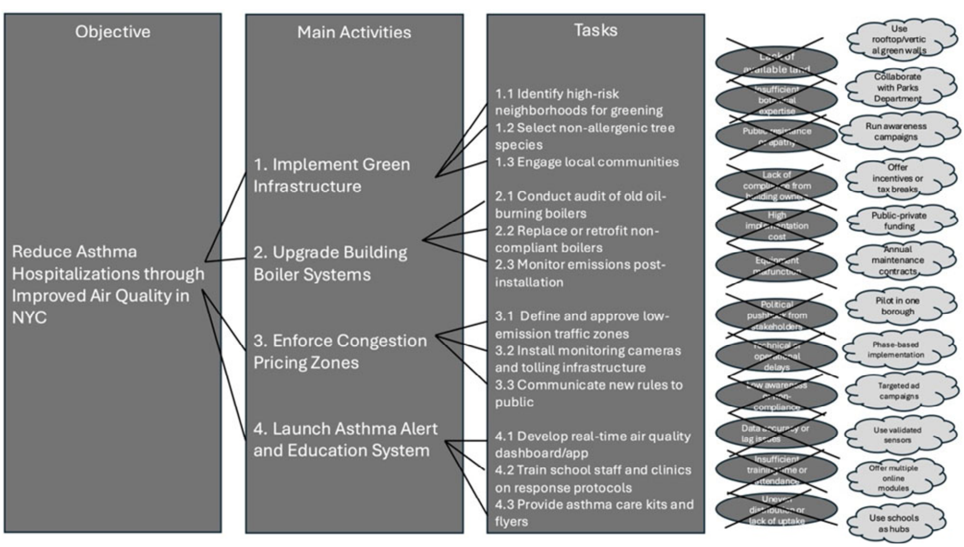
For green infrastructure, the main risks were limited land availability in dense boroughs and community resistance to land repurpose. Proposed mitigations included rooftop gardens, vertical green walls, and partnerships with private developers.

Figure 13 Pareto chart of asthma hospitalisation



Source: Our own (Appendix, Notebook S4; input: NYCCAS-data.csv)

Figure 14 PDPC for anticipated implementation barriers and contingency measures across proposed air quality interventions

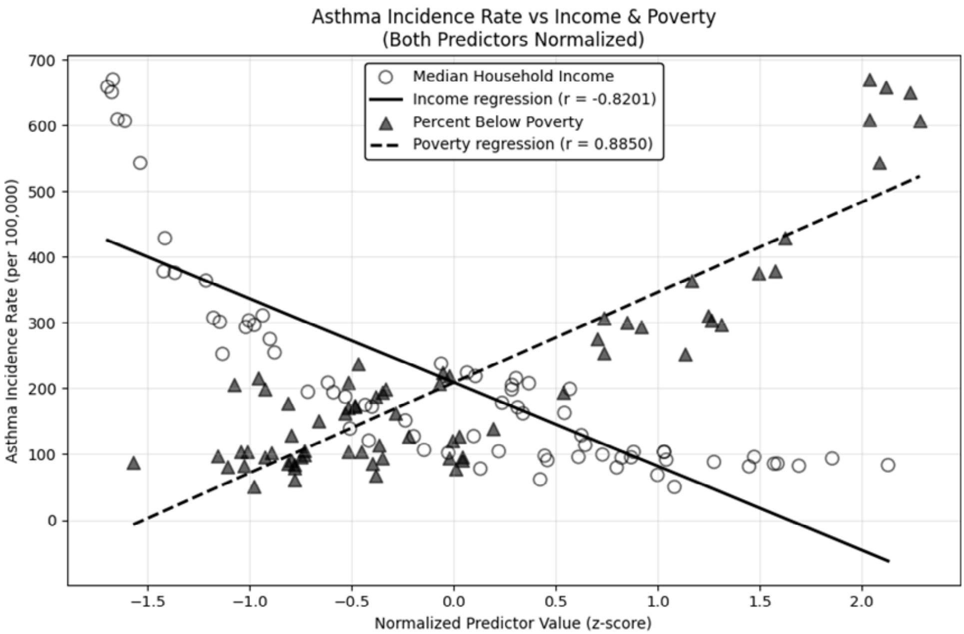


Source: Our own

For boiler modernisation, high retrofit costs, delayed regulatory compliance, and property-owner resistance were the primary barriers. Contingency measures included government subsidies, phased compliance deadlines, and stricter inspection.

For congestion pricing and traffic controls, political opposition, public concern about commuting costs, and infrastructure readiness were identified as the main risks. These can be addressed through phased borough-level implementation, investment in public transport alternatives, and targeted public communication.

Figure 15 Combined scatter plot of poverty rate (%) and median household income against asthma incidence rate per 100,000 residents across New York City boroughs; both predictors normalised (poverty r = 0.885; income r = –0.820)



Source: Our own (Appendix, Notebook S5; input: merged_asthma_data.csv)

For the asthma alert and education system, risks included limited healthcare outreach capacity, digital accessibility gaps in low-income communities, and inconsistent public participation. Proposed responses included multilingual campaigns, school and community health centre partnerships, and local distribution hubs for prevention resources.

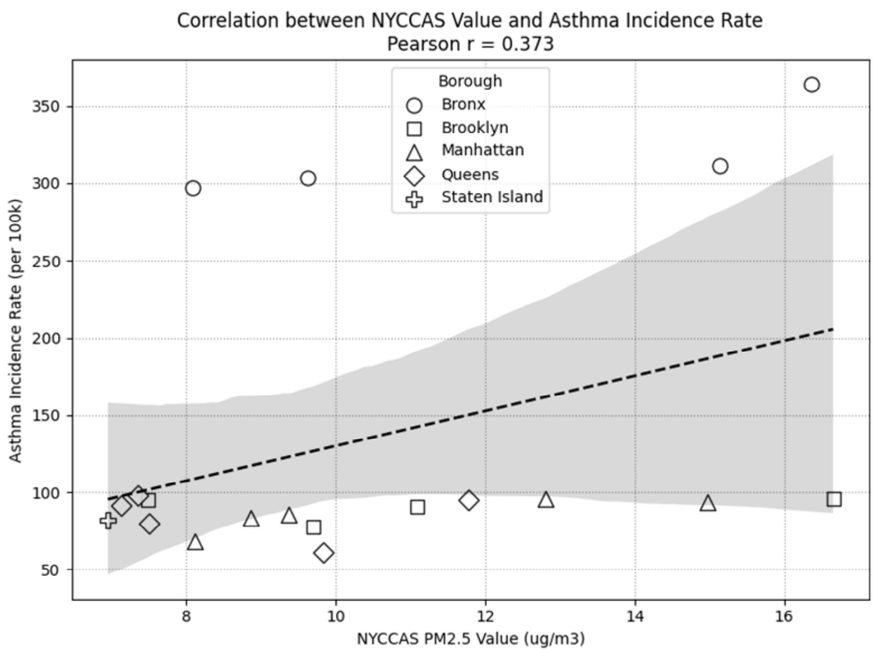
5.7 Synthesis and proposed interventions

Figure 15 plots poverty rate and median household income against asthma incidence on a normalised scale, putting both predictors in the same view. The two regression lines cross near zero. This is the point where borough-level asthma rates sit around the mean before diverging at the extremes. Poverty rate returns a Pearson r of 0.885: at the high end of the poverty distribution, asthma rates cluster between 600 and 670 per 100,000; at the low end, rates fall below 100. The income relationship runs in reverse at r = –0.820, with the lowest-income boroughs in the upper-left of the chart and the highest-income boroughs drifting to the lower-right where rates bottom out below 150 per 100,000.

The two regression lines form an X across the chart. Poverty and incomes are measuring the same underlying condition from different directions, and both point to the same boroughs carrying a disproportionate asthma burden. That consistency across two independent measures makes a stronger case than either correlation alone. Where people can afford to live in New York City is closely tied to how often they end up hospitalised for asthma.

The overall $PM_{2.5}$ correlation chart (Figure 16) returned a Pearson r of 0.369 across all boroughs combined, a moderate positive association. The borough-disaggregated chart (Figure 17) shows why the aggregate is modest. The Bronx borough with $PM_{2.5}$ values reaching 16–17 μg/m^3 in certain years and asthma incidence exceeding 360 per 100,000, pulls the correlation up substantially. Brooklyn, Manhattan, Queens, and Staten Island cluster between 7 and 13 μg/m^3 with correspondingly lower asthma rates, generally below 100 per 100,000. $PM_{2.5}$ alone does not explain borough-level differences. The socio-economic factors in Figures 15 and 16 appear to act as confounding or compounding variables.

Figure 16 Scatter plot of $PM_{2.5}$ concentrations and asthma incidence rate per 100,000 residents across all New York City boroughs combined (Pearson r = 0.369)

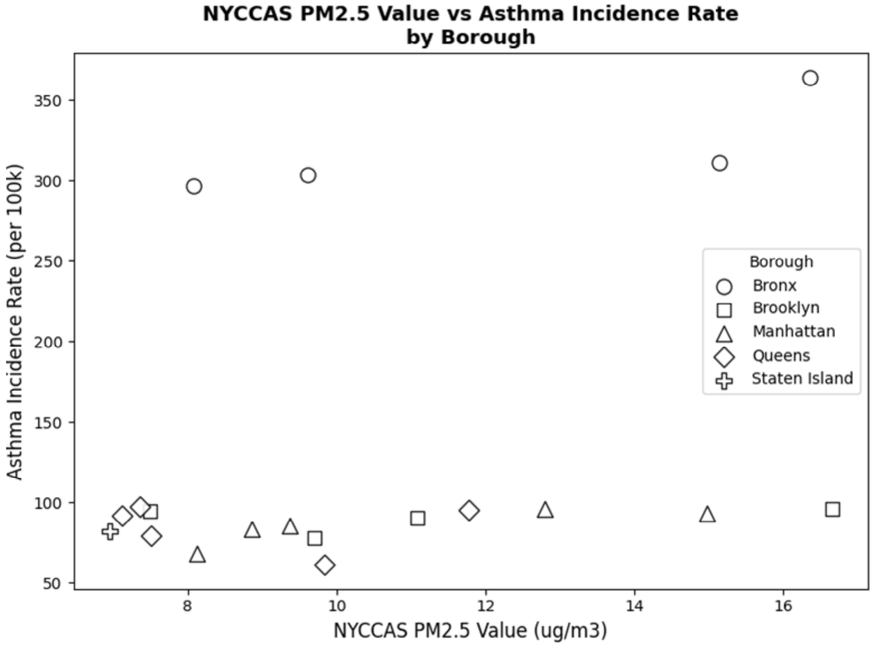


Source: Our own (Appendix, Notebook S5; input: merged_asthma_data.csv)

The poverty control chart by borough (Figure 18) adds longitudinal context. The Bronx maintained the highest poverty rates across the study period, peaking near 29–30% in the early 2010s before declining to approximately 23% by 2024. Those structural conditions, persistent across 15 years, explain why Bronx continues to appear as an outlier in the asthma incidence chart even as citywide pollution levels have fallen.

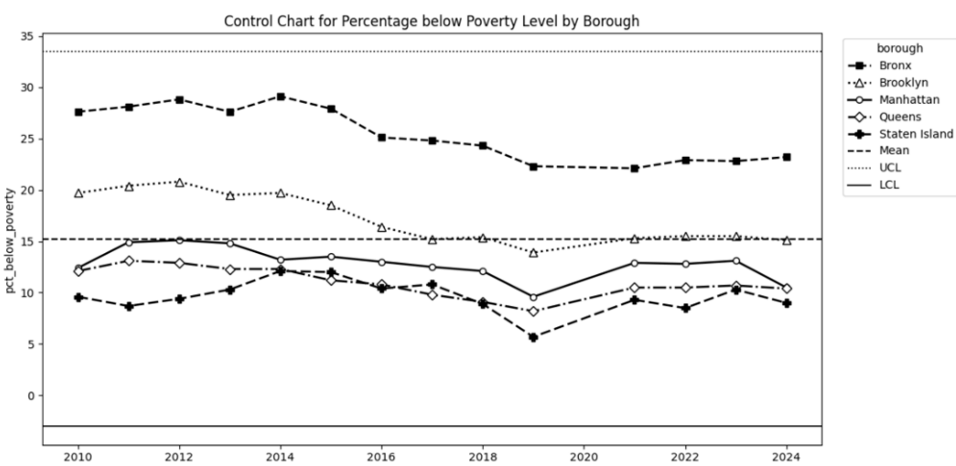
The analyses point toward a consistent finding: elevated pollution from traffic and industrial emissions is strongly associated with higher asthma hospitalisations, and that association is concentrated geographically and amplified by socio-economic vulnerability. Broad, unfocused strategies will underperform targeted ones.

Figure 17 Borough-disaggregated scatter plot of $PM_{2.5}$ concentrations and asthma incidence rate per 100,000 residents, highlighting the Bronx as a high-exposure, high-incidence outlier



Source: Our own (Appendix, Notebook S5; input: merged_asthma_data.csv)

Figure 18 Longitudinal control chart of poverty rate (%) by borough, New York City, 2010–2024



Source: Our own (Appendix, Notebook S6; input: merged_asthma_data.csv)

Several interventions follow from the evidence. Environmental regulation and health surveillance should concentrate in the boroughs that consistently appear as statistical outliers. Bronx appears on the top of the list, given its position across asthma, income, poverty, and $PM_{2.5}$ indicators. Emissions reduction should address both vehicle traffic and industrial sources through congestion pricing, electric vehicle incentives, updated industrial standards, and onsite monitoring. Environmental monitoring infrastructure needs expanded sensor coverage, real-time tracking, and integration with public health alert platforms. Public awareness initiatives should target indoor air quality and transport behaviour, with community organisations continuing to promote asthma trigger awareness. Policy frameworks should align more closely with WHO air quality guidelines while providing incentives for organisations that reduce emissions. Finally, preventive asthma programs and mobile health clinics deployed during pollution spikes should be available in the Bronx and high-poverty areas of Brooklyn, to reduce emergency admissions and build some resilience into communities that have been absorbing this burden the longest.

6 Discussion

The patterns in this data are consistent with what the epidemiological literature has been showing for years: ambient air pollution drives asthma hospitalisations in New York City, and that burden is not shared equally. $PM_{2.5}$ and traffic-related emissions are the primary environmental drivers. Additionally, residency location in the city, the income level, housing quality, and access to healthcare also determine how badly pollution affects you.

6.1 Air pollution as a driver of asthma morbidity

The hospitalisation rates we observed likely reflect two distinct mechanisms operating simultaneously. The first is cumulative developmental exposure. Zanobetti et al. (2024), drawing on the large multi-cohort ECHO CREW Consortium, found that early-life expo-sure to air pollution was associated with cumulative asthma across childhood, implying the damage from living in a high-pollution environment that begins before the school age and compounds over time. Patel et al. (2011) found that traffic density and stationary pollution sources in New York City were associated with wheeze, asthma, and elevated immunoglobulin E levels from birth through age five. By the time a child from the Bronx borough is hospitalised for asthma, years of exposure have already shaped their respiratory physiology.

Further, pollution spikes trigger rapid clinical deterioration, particularly in already vulnerable patients. De et al. (2024) found that pollution episodes were significantly associated with respiratory events and pain crises in children with sickle cell disease in New York City. Madani and Carpenter (2023), in a statewide analysis, found emergency room visits for respiratory disease meaningfully associated with elevated pollution concentrations, alongside poverty and smoking as compounding factors. Smoking is not only a compounding risk factor but a modifiable one. Yellin et al. (2026) examined smoking cessation efforts among patients with asthma and COPD, underscoring cessation support as a point of clinical intervention for this vulnerable population. The hospitalisations in our dataset most certainly represent the severe end of a much larger spectrum of pollution-driven respiratory events occurring across the city.

6.2 Socio-economic inequity and environmental vulnerability

Pollution exposure in New York City is not uniformly distributed, and neither is the hospitalisation burden. Pitiranggon et al. (2025) used land-use regression modelling of $PM_{2.5}$ constituents to identify systematic inequities in exposure across the city to infer about some communities carry a disproportionately heavy load. At the national level, Tessum et al. (2021) demonstrated that $PM_{2.5}$ pollution sources disproportionately and systematically affect people of colour across the USA. New York City, diverse and residentially segregated, is not an exception.

Poverty and pollution are intertwined. Sharma et al. (2023) found that neighbourhood violence and socio-economic deprivation modified the association between acute air pollution and childhood asthma hospitalisations in New York City meaning the same pollution event produces worse outcomes in a deprived neighbourhood than in an affluent one. Munoz-Pizza et al. (2020) reviewing international literature, identified socio-economic status as a consistent effect modifier in the relationship between air pollution and children's respiratory health. The disparities we observed are not explained by pollution alone. They emerge from the intersection of environmental exposure and structural disadvantage, a finding reinforced by Kannoth et al. (2025), who found that chronic air pollution exposure interacted with neighbourhood environmental vulnerability to worsen outcomes among hospitalised New York City residents during the COVID-19 period. Cultural and demographic factors also moderate this relationship. Opoola et al. (2026), in a pilot study of African American children in Harlem, examined how acculturation shapes asthma outcomes, adding a further dimension to the socio-economic and structural disparities already documented in the borough.

6.3 Childhood asthma, physical health, and indoor environments

Outdoor air quality is only part of the picture. Cochran et al. (2023) found associations between fungal diversity in New York City homes and asthma morbidity among school-age children, pointing to indoor environments as a distinct and underappreciated exposure pathway, particularly in dense urban housing where remediation is expensive and often deferred. This is consistent with Kimpo et al. (2026), who found that visible mould in the home was associated with missed school days for asthma among children living in neighbourhoods with higher asthma prevalence, linking indoor fungal exposure directly to an academic and functional outcome. Feasibility work by George et al. (2026) further suggests that indoor air quality monitoring is achievable in low-income housing among Black adults with uncontrolled asthma, pointing to a practical pathway for surveillance in exactly the housing stock where remediation is hardest to reach.

The downstream consequences of poorly controlled asthma compound over time. D'Agostino et al. (2023) found a longitudinal association between asthma severity and reduced physical fitness among New York City public school youth. These researchers found that poorly controlled asthma contributes to physical deconditioning, which may itself worsen long-term respiratory outcomes. Hardell et al. (2023), focusing on the Bronx borough, found that pollutant exposure was associated with greater length of hospital stay among children admitted for asthma. Pollution does appear to make each episode more severe. Vulnerable paediatric subgroups warrant attention. Islam et al. (2026) found elevated asthma and atopy prevalence among perinatally HIV-exposed but uninfected children in Brooklyn, indicating that immune and developmental history may compound the indoor and outdoor exposures already discussed.

6.4 The COVID-19 pandemic intervention

The pandemic disrupted asthma epidemiology in ways that complicated interpretation of any time-series analysis, spanning 2020–2022. Thanik et al. (2023) documented a significant decline in paediatric asthma-related healthcare utilisation in New York City during the pandemic, attributable to reduced outdoor activity, altered pollution patterns from lower traffic volumes, and reluctance to seek in-person care. Skochko et al. (2024) found notable shifts in the spatial distribution of asthma burden before and during the pandemic. McCuskee et al. (2025), examining a large urban health system from 2018 to 2023, found that viral transmission and environmental exposures are increasingly important determinants of paediatric asthma severity in the post-pandemic period, suggesting the relationship between infection and environmental factors is still evolving.

The international picture adds further caution. Choi and Hong (2025) argued against reading the apparent decline in childhood asthma prevalence in 2022 as a genuine epidemiological shift, attributing it instead to pandemic-related disruptions to diagnosis and healthcare access. Stadhouders et al. (2025) documented suppressed exacerbation rates during lockdowns followed by a rebound as social mixing resumed. Panichaporn et al. (2025) found that asthma control levels in children and adolescents changed across the pandemic period. Taken together, these studies support treating 2020–2022 hospitalisation data with caution rather than as a straightforward signal.

6.5 Policy and managerial implications

Hospitalisations related to asthma in New York City appear to result from the combined effects of cumulative developmental exposures and acute fluctuations in air pollution, with substantial geographic and socio-economic disparities across boroughs. These patterns underscore the limitations of relying on citywide averages and support the development of targeted, neighbourhood specific interventions. Policy efforts should prioritise reductions in early-life environmental exposures in high burden communities, particularly in the Bronx, Southern Brooklyn, and Central Queens, where elevated ozone-related mortality and fine particulate matter concentrations continue to pose significant public health risks. Strengthened enforcement of building emission standards and targeted traffic management strategies are also warranted in communities experiencing disproportionate environmental burdens.

Effective management requires greater coordination between public health, environmental, housing, and healthcare institutions. Such collaboration should integrate outdoor pollution mitigation with interventions addressing indoor air quality, fungal exposures, and housing conditions in densely populated neighbourhoods. Because hospitalisation records alone may underestimate the overall burden of asthma, healthcare administrators should integrate emergency department, primary care, and environmental data using advanced analytical methods to strengthen risk prediction, resource allocation, and equity adjusted performance assessment. Post pandemic changes in healthcare utilisation and evolving interactions among respiratory viruses and environmental exposures further require cautious interpretation of declining hospitalisation rates and more comprehensive surveillance approaches.

6.6 Limitations of the research and future research directions

This study has several limitations that provide important directions for future research. Hospitalisation data may underestimate the overall burden of asthma because many acute episodes are treated in emergency departments or primary care settings, while others remain untreated due to insurance limitations, healthcare access barriers, or reluctance to seek medical care. In addition, the ecological nature of the data limits individual level causal inference. Borough level measures of pollution and hospitalisation may conceal substantial variation in individual exposures, household conditions, and neighbourhood microenvironments.

Future research should employ integrated and methodologically advanced approaches to address these limitations. Linking individual electronic health records with high resolution spatial and temporal pollution data could improve exposure assessment and strengthen causal inference. Advanced analytical approaches, including deep learning models combining clinical and environmental data, may further enhance disease prediction and detection accuracy (Migdady et al., 2025). Qualitative research examining healthcare seeking behaviour and barriers to access could provide additional insights into disparities in service utilisation. Longitudinal studies following individuals from early-life through childhood could clarify cumulative exposure pathways and developmental effects. Finally, integrating indoor and outdoor air quality measures would provide a more comprehensive assessment of environmental risks, particularly among populations living in densely populated urban environments.

7 Conclusions

Ambient air pollution, $PM_{2.5}$ and NO_2, drives asthma hospitalisations across New York City's boroughs, and the Bronx borough has borne the largest share of that burden throughout the 2010–2024 study period. The quality analytics tools such as control charts, Pareto analysis, affinity diagrams, interrelationship digraphs, and the PDPC framework clarify that the pollution burden is neither spatially uniform nor purely environmental. Poverty rate ($r = 0.885$) and median household income ($r = –0.820$) were the strongest predictors of asthma incidence in the data, stronger than the aggregate pollutant correlations.

Air quality across the city has improved since roughly 2014. $PM_{2.5}$ and NO_2 trends are both down. But that progress has not reached the Bronx borough in any proportionate way. Asthma incidence in Bronx remains well above the citywide means, sustained by elevated pollution exposure, persistent poverty, high housing density, and limited access to preventive care. Winter brings predictable spikes in both $PM_{2.5}$ and NO_2 that track directly with peaks in asthma admissions. The 2023 Canadian wildfire smoke episode showed that communities already carrying high pollution burdens are also the most exposed communities when transboundary events add to the load.

The pandemic briefly broke these patterns. Lower traffic cut emissions leading to lower hospitalisations, but post-pandemic trends point to a partial rebound in asthma severity. That rebound is a reminder that improvement during an anomalous period is not the same as structural change.

The interventions that follow from this analysis are geographic and specific: concentrating environmental regulation and health surveillance in Bronx and other high-burden boroughs, strengthen real-time pollution monitoring, align policy with WHO air quality guidelines, and deploy community-level healthcare support during pollution spike events. The broader point is harder to act on but more important fact is that pollution and poverty are not independent problems in New York City. Any strategy that addresses one without the other will suboptimise the impact.

Declarations

All authors declare that they have no conflicts of interest.

References

Alkiayat, M. (2021) 'A practical guide to creating a Pareto chart as an quality improvement tool', *Global Journal on Quality and Safety in Healthcare*, Vol. 4, No. 2, pp.83–84, DOI: 10.36401/JQSH-21-X1.

Buss, J., Carpenter, K.P., Krings, J.G., Razi, S., Keller, M., Lyons, P.G., Castro, M., Goss, C.W. and Abbott, M.R. (2026) 'Association between noninvasive positive pressure ventilation use and clinical outcomes during a severe asthma exacerbation: a cohort study', *Critical Care Medicine*, Vol. 54, No. 3, pp.519–531, DOI: 10.1097/CCM.0000000000007025.

Chen, K., Ma, Y., Bell, M.L. and Yang, W. (2023) 'Canadian wildfire smoke and asthma syndrome emergency department visits in New York City', *JAMA*, Vol. 330, No. 14, p.1385, DOI: 10.1001/jama.2023.18768

Choi, E.J. and Hong, S-J. (2025) 'The trend of childhood asthma prevalence decreased in 2022: true or not?', *Allergy, Asthma & Immunology Research*, Vol. 17, No. 3, p.285, DOI: 10.4168/aair.2025.17.3.285.

Cochran, S.J., Acosta, L., Divjan, A., Lemons, A.R., Rundle, A.G., Miller, R.L., Sobek, E., Green, B.J., Perzanowski, M.S. and Dannemiller, K.C. (2022) 'Spring is associated with increased total and allergenic fungal concentrations in house dust from a pediatric asthma cohort in New York City', *Building and Environment*, Vol. 226, p.109711, DOI: 10.1016/j.buildenv.2022.109711

Cochran, S.J., Acosta, L., Divjan, A., Lemons, A.R., Rundle, A.G., Miller, R.L., Sobek, E., Green, B.J., Perzanowski, M.S. and Dannemiller, K.C. (2023) 'Fungal diversity in homes and asthma morbidity among school-age children in New York City', *Environmental Research*, Vol. 239, p.117296, DOI: 10.1016/j.envres.2023.117296.

D'Agostino, E.M., Zhang, S., Day, S.E., Konty, K.J., Armstrong, S., Skinner, A. and Neshteruk, C.D. (2023) 'The longitudinal association between asthma severity and physical fitness among New York City public school youth', *Preventive Medicine*, Vol. 170, p.107486, DOI: 10.1016/j.ypmed.2023.107486.

De, A., Jung, K.H., Davis, H., Siddiqui, A., Kattan, M., Quinn, J., Rundle, A., Green, N.S. and Lovinsky-Desir, S. (2024) 'Effects of air pollution on respiratory events and pain crises among children with sickle cell disease in New York City', *Annals of the American Thoracic Society*, Vol. 21, No. 12, pp.1733–1741, DOI: 10.1513/AnnalsATS.202310-860OC.

Deepu, T.S. and Ravi, V. (2026) 'Prioritisation of critical operational factors and designing a decision-making model for attaining supply chain resilience through digitalisation', *International Journal of Business Information Systems*, Vol. 51, No. 1, pp.99–127, DOI: 10.1504/IJBIS.2026.150760.

Espejo, D., Plaza, V., Quirce, S., Trigueros, J.A. and Muñoz, X. (2025) 'Influence of out-door air pollutants on asthma: a narrative review', *Open Respiratory Archives*, Vol. 7, No. 3, p.100448, DOI: 10.1016/j.opresp.2025.100448.

Flores, N.M., Walters, S., Do, V., Hernández, D., Perzanowski, M.S., Rudolph, K.E., Northrop, A.J., Spira-Cohen, A., Lovinsky-Desir, S. and Casey, J.A. (2026) 'Electrical power outages and asthma-related emergency department visits in New York City, 2018–2022', *Epidemiology*, Vol. 37, No. 4, pp.544–552, DOI: 10.1097/EDE.0000000000001991.

George, M., Khurana, R.K., Bruzzese, J-M., Zeremski, M., DiMango, E., Naseem, A., Cassells, A., Mak, E., Benda, N., Tobin, J.N. and Chillrud, S. (2026) 'Assessing the feasibility of indoor air quality monitoring in low-income housing among Black adults with uncontrolled asthma in New York City', *BMC Public Health*, DOI: 10.1186/s12889-026-28295-z.

Ghozlan, B.R. and Shanab, E.A.A. (2025) 'The contribution of data visualisation tools towards better customer relationship management', *International Journal of Business Information Systems*, Vol. 48, No. 1, pp.72–98, DOI: 10.1504/IJBIS.2025.144076.

Goldberg, R., Spira-Cohen, A., Pitiranggon, M., Johnson, S. and Ito, K. (2025) 'Changes in the short-term relationship between air pollution and mortality in New York City, 1990–2019', *Environmental Health*, Vol. 24, No. 1, p.37, DOI: 10.1186/s12940-025-01171-w.

Hardell, J., Silver, E.J., Kavouras, I., Lee, D.S. and Gross, E. (2023) 'Childhood asthma in the Bronx, NY; the impact of pollutants on length of hospital stay', *Journal of Asthma*, Vol. 60, No. 12, pp.2160–2169, DOI: 10.1080/02770903.2023.2225607.

Humphrey, J.L., Kinnee, E.J., Robinson, L.F. and Clougherty, J.E. (2024) 'Disentangling impacts of multiple pollutants on acute cardiovascular events in New York City: a case-crossover analysis', *Environmental Research*, Vol. 242, p.117758, DOI: 10.1016/j.envres.2023.117758.

Islam, S., Kohlhoff, S., Joks, R. and Smith-Norowitz, T. (2026) 'Prevalence of asthma and atopy in perinatally HIV-exposed but uninfected children in Brooklyn, New York', *Journal of Allergy and Clinical Immunology*, Vol. 157, No. 2, Suppl., p.AB243, DOI: 10.1016/j.jaci.2025.12.724.

Johnson, S., Haney, J., Cairone, L., Huskey, C. and Kheirbek, I. (2020) 'Assessing air quality and public health benefits of New York City's climate action plans', *Environmental Science & Technology*, Vol. 54, No. 16, pp.9804–9813, DOI: 10.1021/acs.est.0c00694.

Kannoth, S., Zhang, C., Shafiq, M., Albrecht, S.S., Azan, A., Chambers, E.C., Qian, M., Sheffield, P.E., Thompson, A., Woo Baidal, J.A., Lovinsky-Desir, S. and Stingone, J.A. (2025) 'The relationship between chronic air pollution exposure, neighborhood environmental vulnerability, and adverse COVID-19 morbidities among hospitalized New York City residents', *Environment International*, Vol. 202, p.109660, DOI: 10.1016/j.envint.2025.109660.

Kimpo, J., Acosta, L., Dannemiller, K.C., Lovinsky-Desir, S., Miller, R.L., Rundle, A.G. and Perzanowski, M. (2026) 'C25-08 visible mold in homes is associated with missed school days for asthma among children living in higher asthma prevalence neighborhoods of New York City', *American Journal of Respiratory and Critical Care Medicine*, Vol. 212, No. Suppl. 1, p.aamag162.538, DOI: 10.1093/ajrccm/aamag162.538.

Lau, K., Guo, J., Miao, Y., Ross, Z., Riley, K.W., Wang, S., Herbstman, J. and Perera, F. (2024) 'Major air pollution and climate policies in NYC and trends in NYC air quality 1998–2021', *Frontiers in Public Health*, Vol. 12, p.1474534, DOI: 10.3389/fpubh.2024.1474534.

Madani, N.A. and Carpenter, D.O. (2023) 'Patterns of emergency room visits for respiratory diseases in New York State in relation to air pollution, poverty and smoking', *International Journal of Environmental Research and Public Health*, Vol. 20, No. 4, p.3267, DOI: 10.3390/ijerph20043267.

Mananga, E.S., Lopez, E., Diop, A., Dongomale, P.J. and Diane, F. (2023) 'The impact of air pollution on health in New York City', *Journal of Public Health Research*, Vol. 12, No. 4, p.22799036231205870, DOI: 10.1177/22799036231205870.

McCuskee, S., Kowalski, R., Mei, C., Zebrowski, A., Patel, A., Zajac, L., Zikry, H., Ayyash, A. and DeFelice, N. (2025) 'Increasing severity, importance of viral transmission, and environmental exposures for pediatric asthma in a large urban health system, 2018–2023', *medRxiv*, DOI: 10.1101/2025.02.01.25321524.

Migdady, A., Alaiad, A., Ra, N.A. and Khatib, E.M.A. (2025) 'EfficientNet deep learning model for pneumothorax disease detection in chest X-ray images', *International Journal of Business Information Systems*, Vol. 50, No. 1, pp.1–21, DOI: 10.1504/IJBIS.2025.148498.

Munoz-Pizza, D.M., Villada-Canela, M., Reyna, M.A., Texcalac-Sangrador, J.L. and Osornio-Vargas, Á.R. (2020) 'Air pollution and children's respiratory health: a scoping review of socioeconomic status as an effect modifier', *International Journal of Public Health*, Vol. 65, No. 5, pp.649–660, DOI: 10.1007/s00038-020-01378-3.

Opoola, Z.A., Jung, K., Aubey, B. and Lovinsky-Desir, S. (2026) 'A60-05 acculturation and asthma outcomes among African American children in Harlem, New York City: a pilot study', *American Journal of Respiratory and Critical Care Medicine*, Vol. 212, No. Suppl. 1, p.aamag162.353, DOI: 10.1093/ajrccm/aamag162.353.

Panichaporn, W., Boonard, K., Kanchanapoomi, K., Srisuwatchari, W., Jirapongsananuruk, O., Visitsunthorn, N. and Pacharn, P. (2025) 'Level of asthma control in children and adolescents before and during the COVID-19 pandemic', *Asian Pacific Journal of Allergy and Immunology*, Vol. 43, No. 2, pp.312–319, DOI: 10.12932/AP-180124-1770.

Patel, M.M., Quinn, J.W., Jung, K.H., Hoepner, L., Diaz, D., Perzanowski, M., Rundle, A., Kinney, P.L., Perera, F.P. and Miller, R.L. (2011) 'Traffic density and stationary sources of air pollution associated with wheeze, asthma, and immunoglobulin E from birth to age 5 years among New York City children', *Environmental Research*, Vol. 111, No. 8, pp.1222–1229, DOI: 10.1016/j.envres.2011.08.004.

Pitiranggon, M., Johnson, S., Spira-Cohen, A., Eisl, H. and Ito, K. (2025) 'Determining sources of air pollution exposure inequity in New York City through land-use regression modeling of $PM_{2.5}$ constituents', *Pollutants*, Vol. 5, No. 1, p.2, DOI: 10.3390/pollutants5010002.

Romaszko-Wojtowicz, A., Dragańska, E., Doboszyńska, A. and Glińska-Lewczuk, K. (2025) 'Impact of seasonal biometeorological conditions and particulate matter on asthma and COPD hospital admissions', *Scientific Reports*, Vol. 15, No. 1, p.450, DOI: 10.1038/s41598-024-84739-9.

Sharma, R., Humphrey, J.L., Frueh, L., Kinnee, E.J., Sheffield, P.E. and Clougherty, J.E. (2023) 'Neighborhood violence and socioeconomic deprivation influence associations between acute air pollution and temperature on childhood asthma in New York City', *Environmental Research*, Vol. 231, p.116235, DOI: 10.1016/j.envres.2023.116235.

Shukla, K., Seppanen, C., Naess, B., Chang, C., Cooley, D., Maier, A., Divita, F., Pitiranggon, M., Johnson, S., Ito, K. and Arunachalam, S. (2022) 'ZIP code-level estimation of air quality and health risk due to particulate matter pollution in New York City', *Environmental Science & Technology*, Vol. 56, No. 11, pp.7119–7130, DOI: 10.1021/acs.est.1c07325.

Skochko, S., Nguyen, T., Mack, S., Turcotte, B., Adler, C., Rosenberg, E.S., Joseph, C., Siag, L., Dubuisson, A. and Wagner, V.L. (2024) 'Asthma hot spots in New York before and during the COVID-19 pandemic', *Preventing Chronic Disease*, Vol. 21, p.240059, DOI: 10.5888/pcd21.240059.

Stadhouders, L., Hoogteijling, E.M., Duijts, L. and Lebon, A. (2025) 'Trends in asthma exacerbations in children before, during, and after the COVID-19 pandemic', *Epidemiology and Infection*, Vol. 153, p.e138, DOI: 10.1017/S0950268825100800.

Tessum, C.W., Paolella, D.A., Chambliss, S.E., Apte, J.S., Hill, J.D. and Marshall, J.D. (2021) '$PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States', *Science Advances*, Vol. 7, No. 18, p.eabf4491, DOI: 10.1126/sciadv.abf4491.

Thanik, E., Harada, K., Garland, E., Bixby, M., Bhatia, J., Lopez, R., Galvez, S., Dayanov, E., Vemuri, K., Bush, D. and DeFelice, N.B. (2023) 'Impact of COVID-19 on pediatric asthma-related healthcare utilization in New York City: a community-based study', *BMC Pediatrics*, Vol. 23, No. 1, p.41, DOI: 10.1186/s12887-023-03845-1.

Thomas, A. and Suresh, M. (2025) 'Development of a sustainability assessment framework for hospitals using multi-grade fuzzy', *International Journal of Business Information Systems*, Vol. 50, No. 1, pp.113–135, DOI: 10.1504/IJBIS.2025.148501.

Thurston, G., Yu, W. and Luglio, D. (2023) 'An evaluation of the asthma impact of the June 2023 New York City wildfire air pollution episode', *American Journal of Respiratory and Critical Care Medicine*, Vol. 208, No. 8, pp.898–900, DOI: 10.1164/rccm.202306-1073LE.

Yellin, S., Rauhut, M., Kutscher, E. and Anselm, E. (2026) 'Smoking cessation efforts for patients with asthma and COPD', *medRxiv*, DOI: 10.64898/2026.02.14.26345148.

Zanobetti, A., Ryan, P.H., Coull, B.A., Luttmann-Gibson, H., Datta, S., Blossom, J., Brokamp, C., Lothrop, N., Miller, R.L., Beamer, P.I., Visness, C.M., Andrews, H., Bacharier, L.B., Hartert, T., Johnson, C.C., Ownby, D.R., Khurana Hershey, G.K., Joseph, C.L.M., Mendonça, E.A., Jackson, D.J., Zoratti, E.M., Wright, A.L., Martinez, F.D., Seroogy, C.M., Ramratnam, S.K., Calatroni, A., Gern, J.E. and Gold, D.R. (2024) 'Early-life exposure to air pollution and childhood asthma cumulative incidence in the ECHO CREW Consortium', *JAMA Network Open*, Vol. 7, No. 2, p.e240535, DOI: 10.1001/jamanetworkopen.2024.0535.

Zhou, X., Sampath, V. and Nadeau, K.C. (2024) 'Effect of air pollution on asthma', *Annals of Allergy, Asthma & Immunology*, Vol. 132, No. 4, pp.426–432, DOI: 10.1016/j.anai.2024.01.017.

Appendix

Table A1 Summary of data sources and temporal coverage

Data source	*Variables used*	*Spatial level*	*Time period*	*Original format*
New York City Community Air Survey (NYCCAS)	$PM_{2.5}$ monthly observations and site metadata	Site level, aggregated to borough	October 2019–February 2026	CSV
NYCCAS air quality analytics files	$PM_{2.5}$, NO_2, and O_3 seasonal and annual concentrations	Yearly and borough-level	Historical NYCCAS series used for control charts and forecasting	CSV/chart outputs
SPARCS Inpatient Data	Asthma hospitalisation cases, ED visits, length of stay, severity, mortality risk, charges, costs, and Medicaid indicators	ZIP3, age group, and borough	2010–2024	CSV
American Community Survey (ACS)	Population, income, poverty, insurance, Medicaid, overcrowding, age, and demographic indicators	Borough/county	2010–2024, excluding 2020 where unavailable	API/CSV
Integrated asthma, ACS, and NYC-CAS analysis data	Asthma incidence rate, pollutant values, and socio-economic predictors	Borough-year	2010–2024 subject to source overlap	CSV/visual outputs

Source: Our own

Table A2 Notebook table

Notebook	*File*	*Purpose*
Notebook S1	S1-download-NYCCAS-data.ipynb	Downloads monthly NYCCAS CSV files, merges site-level observations, attaches station metadata, assigns boroughs, and creates seasonal and annual air quality aggregates.
Notebook S2	S2-acs-yearly-socioeconomic-indicators.ipynb	Retrieves annual ACS socio-economic profile indicators for New York City counties and creates a borough-level indicator dataset for 2010–2024.
Notebook S3	S3-sparcs-aggregate.ipynb	Extracts asthma-related SPARCS inpatient discharge records, normalises key variables, computes hospitalisation indicators, and aggregates results by ZIP3, age group, year, and borough.
Notebook S4	S4-nyccas-control-charts-forecasting.ipynb	Generates NYCCAS $PM_{2.5}$, NO_2, and O_3 seasonal control charts and compares short-term forecasting models.
Notebook S5	S5-borough-asthma-incidence-vs-pm25.ipynb.ipynb	Calculates borough-level asthma incidence rates and evaluates relationships between asthma incidence and $PM_{2.5}$ exposure.
Notebook S6	S6-acs-poverty-income-asthma-correlation.ipynb	Analyses poverty, income, insurance, and related ACS indicators against borough-level asthma incidence.
Notebook S7	S7-sparcs-pareto-analysis.ipynb	Performs SPARCS Pareto analysis by hospitalisation reason group and creates ranked cumulative percentage outputs.

Notebook S1: S1-download-NYCCAS-data.ipynb

Parameter	*Detail*
Data source	NYCCAS GitHub raw monthly pollution CSV archive and station metadata from station-new.csv.
Coverage	Monthly observations from October 2019 through February 2026.
Geographic unit	NYCCAS monitoring site level, mapped and aggregated to New York City boroughs.
Key variables retrieved	Observation timestamp, site identifier/name, $PM_{2.5}$ value, month, year, season, and borough.
Processing steps	Creates data_csv/NYCCAS/, downloads monthly files, downloads site-info.csv, merges monthly data, attaches site metadata, maps sites to boroughs, and computes seasonal and annual averages.
Output files	data_csv/NYCCAS_2019-2026.csv; data_csv/NYCCAS/site-info.csv; NYCCAS-Winter/Spring/Summer/Fall files; NYCCAS-Seasonal-Borough.csv; NYCCAS-Yearly-Borough.csv.
Note	Winter is handled as a cross-year season by assigning December observations to the following winter year.

Source: Our own

Table A2 Notebook table (continued)

Notebook S2: S2-acs-yearly-socioeconomic-indicators.ipynb	
Parameter	*Detail*
Data source	U.S. Census Bureau ACS 1-Year Profile API.
Coverage	2010-2024, excluding 2020; 2025 is attempted but skipped when unavailable from the API.
Geographic unit	New York City boroughs represented as county-level ACS records.
Key variables retrieved	Median household income, poverty rate, uninsured rate, Medicaid rate, overcrowded housing, total population, child population share, selected race/ethnicity indicators, and population density.
Variable mapping	Year-aware ACS variable mapping accommodates coding changes before and after 2017.
Output file	data_csv/ACS_socioeconomic_indicators_2010-2024_Data.csv.
Note	The notebook optionally reads an API key from a .env file and combines DP03, DP04, and DP05 profile tables.
Notebook S3: S3-sparcs-aggregate.ipynb	
Parameter	*Detail*
Data source	New York State Department of Health SPARCS inpatient discharge CSV files stored under data_csv/SPARCS/.
Coverage	2010–2024 annual inpatient discharge files.
Geographic unit	ZIP3 and age group, then aggregated to borough-year level for New York City.
Key variables retrieved	Asthma cases, ED visits, average length of stay, average severity, average mortality risk, total charges, total costs, average costs, Medicaid cases, ED rate, and Medicaid rate.
Diagnosis handling	Uses CCS diagnosis descriptions for earlier years and CCSR diagnosis descriptions for later years; asthma cases are selected through asthma diagnosis text matching.
ZIP3 mapping	NYC ZIP3 prefixes are mapped to Manhattan, Bronx, Brooklyn, Queens, and Staten Island; non-NYC ZIP3 records are filtered before borough aggregation.
Output file	data_csv/SPARCS_by_borough_2010-2024.csv.
Notebook S4: S4-nyccas-control-charts-forecasting.ipynb	
Parameter	*Detail*
Input files	data_csv/NYCCAS-PM25.csv; data_csv/NYCCAS-NO2.csv; data_csv/NYCCAS-O3.csv; data_csv/NYCCAS-Seasonal-borough.csv.
Pollutants analysed	$PM_{2.5}$, NO_2, and O_3.
Key analyses	Seasonal control charts, yearly mean plots, borough-level pollutant visualisations, and short-term pollutant forecasting.
Control chart method	Uses grand mean plus or minus three sample standard deviations to calculate UCL and LCL.
Forecasting models	Linear regression, polynomial regression with degree 2, decision tree regression, and exponential smoothing.

Source: Our own

Table A2 Notebook table (continued)

Notebook S4: S4-nyccas-control-charts-forecasting.ipynb	
Parameter	*Detail*
Evaluation metric	Root mean square error (RMSE) over the final three-year test period.
Key outputs	Publication-ready control chart visualisations, borough-level trend plots, forecast comparisons, and RMSE summaries.
Libraries used	pandas, numpy, matplotlib, scikit-learn, and statsmodels.
Notebook S5: S5-borough-asthma-incidence-vs-pm25.ipynb.ipynb	
Parameter	*Detail*
Input files	data_csv/SPARCS_by_borough_2010-2024.csv; data_csv/ACS_socioeconomic_indicators_2010-2024_Data.csv; data_csv/NYCCAS-Yearly-borough.csv.
Geographic unit	Borough-year level.
Key analyses	Borough-level asthma incidence calculation, PM2.5-asthma scatterplots, correlation diagnostics, and asthma incidence control chart.
Incidence formula	(asthma_cases / total_population) × 100,000.
Key outputs	merged_asthma_data.csv; $PM_{2.5}$ versus asthma incidence scatterplot; Pearson correlation; combined asthma incidence control chart by borough.
Output file	data_csv/merged_asthma_data.csv.
Libraries used	pandas, matplotlib, and seaborn.
Notebook S6: S6-acs-poverty-income-asthma-correlation.ipynb	
Parameter	*Detail*
Input files	data_csv/ACS_socioeconomic_indicators_2010-2024_Data.csv; data_csv/merged_asthma_data.csv.
Geographic unit	Borough-year level.
Key analyses	Borough trend plots for ACS socio-economic indicators and correlations between socio-economic predictors and asthma incidence.
Socio-economic predictors	Poverty rate, median household income, insurance-related measures, Medicaid share, population, and related ACS indicators where available.
Key outputs	Poverty and income trend visualisations, asthma incidence comparison plots, and correlation summaries.
Note	The notebook standardises borough and year columns before merging ACS and asthma incidence data.
Libraries used	pandas, numpy, matplotlib, os, and glob.
Notebook S7: S7-sparcs-pareto-analysis.ipynb	
Parameter	*Detail*
Data source	New York State Department of Health SPARCS inpatient discharge files stored under data_csv/SPARCS/.
Coverage	2010–2024, processed in period groups such as 2010–2015, 2016–2020, and 2021–2024 before combining.
Geographic unit	ZIP3, age group, year, and hospitalisation reason group.

Source: Our own

Table A2 Notebook table (continued)

Notebook S7: S7-sparcs-pareto-analysis.ipynb	
Parameter	*Detail*
Key variables retrieved	Hospitalisation reason group, asthma cases, ED visits, length of stay, severity, mortality risk, charges, costs, Medicaid cases, ED rate, and Medicaid rate.
Hospitalisation reason groups	Asthma exacerbation, other respiratory disorders, respiratory failure, respiratory infection, and other grouped respiratory categories identified in the raw SPARCS descriptions.
Key analysis	Pareto chart ranking hospitalisation reason groups by total count with cumulative percentage curve and 80% reference line.
Output files	data_csv/sparcs_asthma_aggregated_2010_2015.csv; data_csv/sparcs_asthma_aggregated_2016_2020.csv; data_csv/sparcs_asthma_aggregated_2021_2024.csv; data_csv/sparcs_asthma_aggregated_2010_2024.csv; data_csv/sparcs_pareto_hospitalisation_reason_groups_2010_2024.csv when saved.
Libraries used	pandas, numpy, matplotlib, glob, os, and re.

Source: Our own

Table A3 Intermediate and processed output files

Output file	*Generating notebook*	*Description*
data_csv/NYCCAS/monthly CSV files	Notebook S1	Downloaded monthly NYCCAS site-level air-quality observations.
data_csv/NYCCAS_2019-2026.csv	Notebook S1	Merged NYCCAS monthly site-level observations for October 2019-February 2026.
data_csv/NYCCAS/site-info.csv	Notebook S1	NYCCAS station metadata used to attach monitoring-site names and locations.
NYCCAS-Seasonal-Borough.csv	Notebook S1	Borough-level seasonal air-quality averages.
NYCCAS-Yearly-Borough.csv	Notebook S1	Borough-level annual air-quality averages.
ACS_socioeconomic_indicators_ 2010-2024_Data.csv	Notebook S2	Annual borough-level ACS demo-graphic and socio-economic indicators.
SPARCS_by_Borough_2010-2024.csv	Notebook S3	Borough-year asthma hospitalisation indicators derived from SPARCS inpatient records.
merged_asthma_data.csv	Notebook S5	Integrated asthma incidence table using SPARCS cases and ACS population.

Source: Our own

Table A3 Intermediate and processed output files (continued)

Output file	*Generating notebook*	*Description*
sparcs_asthma_aggregated_2010_2015.csv; sparcs_asthma_aggregated_2016_2020.csv; sparcs_asthma_aggregated_2021_2024.csv	Notebook S7	Period-level SPARCS asthma aggregation files used for final Pareto analysis.
sparcs_asthma_aggregated_2010_2024.csv	Notebook S7	Combined SPARCS asthma aggregation for the full 2010–2024 period.
sparcs_pareto_hospitalisation_reason_ groups_2010_2024.csv	Notebook S7	Ranked Pareto table of hospitalisation reason groups, counts, percentages, and cumulative percentages when exported.

Source: Our own

Table A4 Quality analytics tools applied

Tool	*Purpose*
Descriptive statistics	Characterise pollutant distributions across boroughs and seasonal periods
Trend analysis	Identify recurring seasonal and longitudinal pollution patterns
Affinity diagram	Organise variables into conceptual clusters linking pollution sources, environmental metrics, and health outcomes
Interrelationship digraph	Visualise cause-and-effect relationships within the environmental health system
Control charts	Evaluate process stability and identify significant variations or outliers in pollution levels
PDPC	Anticipate implementation barriers and map contingency responses for proposed air quality interventions

Source: Our own