

# Mapping Systematically Neglected Neighborhoods: a New York City Case Study

Xinhao Wang

*Xi'an Jiaotong-Liverpool University, Suzhou, China*

**Abstract:** Cities now release large volumes of open government data, which promises greater transparency and accountability. However, publishing data does not by itself tell us whether public services and environmental burdens are shared fairly across a city. This paper presents a reproducible framework for measuring spatial equity using only free and public data. The framework joins three sources at the census tract level: resident service requests from a city 311 system, demographic and income data from the United States Census American Community Survey, and environmental burden indicators from the United States Environmental Protection Agency EJScreen tool. From these we build tract level measures of service responsiveness, environmental burden, and social vulnerability, and we use spatial statistics to find tracts where high burden, slow service, and high vulnerability occur together. We call these tracts systematically neglected. We then apply the service part of the framework to New York City, using one partial year of 311 records from 2025. The run produces tract level maps of complaint volume, average resolution time, unresolved rate, and a combined neglect score, and it flags the twenty five most neglected tracts. The result is informative in an unexpected way. Most flagged tracts fall in Manhattan, which is the wealthiest borough, because the score includes complaint volume and volume tracks the propensity to report rather than true need. This is a live example of the reporting bias that the 311 literature describes, and it shows in practice why raw complaint volume must not be read as a sign of neglect. We discuss how to correct this by down weighting volume and by adding the environmental and demographic layers. We also discuss the removal of EJScreen from the EPA website in 2025 and its volunteer reconstruction, which makes data preservation part of equity work.

The framework can be applied to any city that runs a 311 system.

**Keywords:** Environmental Justice; Spatial Equity; Open Government Data; 311 Service Requests; Reporting Bias; Spatial Autocorrelation; Urban Informatics

## 1. Introduction

Open government data has become a normal part of city administration. Hundreds of cities in North America and beyond now run open data portals where residents and researchers can download records on service requests, transport, housing, and the environment. The promise of this movement is clear. Open data should make government more transparent, support better decisions, and allow the public to hold agencies to account. Yet the existence of a dataset does not answer the question that matters most for fairness. We still need to ask whether public services and environmental burdens are distributed in line with the needs of communities, or whether some neighborhoods receive less and bear more.

This question sits at the center of environmental justice research, which studies how environmental risks and benefits are shared across groups defined by race, ethnicity, and income [1]. A large body of work shows that pollution and other hazards fall more heavily on low income communities and communities of color [1], [2]. A parallel and growing body of work studies the data that cities now collect about themselves. Resident reporting systems, known in the United States as 311 systems, generate millions of records each year and are widely used to measure the condition of a city in close to real time [5], [7]. These two literatures are usually kept apart. Environmental justice work tends to focus on pollution exposure, while urban informatics work tends to focus on service requests. The neighborhoods that need attention most are often those where both problems appear together, that is, high environmental burden and

weak public service response.

This paper proposes a single framework that brings these strands together. The goal is to identify what we call systematically neglected neighborhoods. We define a systematically neglected tract as one that combines three conditions: a high environmental burden, a slow or weak public service response, and a socially vulnerable population. We design the framework so that it relies only on free and public data, and so that it can be reproduced in any city that publishes 311 records. We pay careful attention to the known weaknesses of these data, especially the fact that residents do not report problems at equal rates, which can make a city look responsive in places that simply complain more.

We then test the service part of the framework on New York City. The case study is useful in two ways. It shows that the pipeline runs end to end on real open data, and it shows, in a concrete way, how reporting bias can shape the result if the analyst is not careful. As we report below, a naive neglect score that includes raw complaint volume flags mostly wealthy central tracts. This is not a finding about true neglect. It is a warning about method, and it points to the correction that the full framework provides.

The contribution of this paper is fourfold. First, we describe a reproducible pipeline that joins 311 service requests, Census demographic data, and EPA environmental indicators at the census tract level. Second, we define clear tract level measures and combine them into an equity screen. Third, we apply established spatial statistics to detect clusters of disadvantage. Fourth, we treat reporting bias as a first order concern, and our New York City results show why this matters. We also reflect on a recent event that makes this work timely. In early 2025 the EPA removed its EJScreen tool from public access, and a volunteer coalition reconstructed it so that researchers could continue to use it [15]. This event shows that the data behind equity analysis can be removed for political reasons, and it makes the case that reproducibility and data preservation are themselves equity concerns.

## **2. Related Work**

### **2.1 Environmental Justice and Distributive Equity**

Environmental justice research asks whether environmental harms and benefits are shared

fairly. Mohai, Pellow, and Roberts review two decades of evidence and conclude that exposure to pollution is unequally distributed by race and class [1]. An important methodological point in this literature is that the answer depends on how space is measured. Mohai and Saha show that the choice of spatial unit and the method used to link populations to hazards can change the size and even the direction of measured disparities [2]. This is a warning for any spatial study, including ours. Results must be reported together with the units and the joining method used to produce them.

Later work argues that single hazards understate the problem, because disadvantaged communities often face many burdens at once. Morello-Frosch and colleagues call for the study of cumulative impacts, where social and environmental stressors combine and reinforce each other [3]. The EJScreen tool that we use is built on this idea, since it reports many environmental indicators side by side with demographic indicators. The same logic of distribution applies to benefits, not only burdens. Wolch, Byrne, and Newell review the evidence on urban green space and find that parks and green areas often favor wealthier and whiter neighborhoods, which makes access to green space an environmental justice issue in its own right [4]. Our framework treats burdens and the absence of services in a parallel way, because both shape the lived quality of a neighborhood.

### **2.2 311 Service Requests and Reporting Bias**

311 systems let residents request services and report non emergency problems by phone, web, or mobile app. They have become a major link between residents and city government and a rich source of data for research [5], [7]. Researchers have used these data to classify neighborhoods, to predict socioeconomic conditions, and to study how residents engage with government [6], [7], [8]. The central caution in this literature is that 311 data record who reports a problem, not where problems actually exist. Kontokosta, Hong, and Korsberg study more than twenty million New York City requests and show that the propensity to complain varies with the socioeconomic and demographic profile of an area, with factors such as English language ability playing a role [5]. O'Brien shows that 311 reporting reflects patterns of civic engagement and territorial behavior, not only objective need [6]. Minkoff

provides a tract level analysis of who contacts the city and confirms that contact rates are socially patterned [8].

Recent work makes this concern sharper. Liu, Bhandaram, and Garg develop a method to estimate reporting delays without external ground truth, by using the rate of duplicate reports about the same incident [9]. Applying their method to more than one hundred thousand reports in New York City and more than nine hundred thousand in Chicago, they find large spatial differences in how quickly incidents are reported, even after controlling for the type of incident, with some neighborhoods reporting about three times faster than others, and these differences track socioeconomic characteristics [9]. The lesson for our framework is direct, and our own results confirm it. We cannot treat the raw volume of complaints as a measure of need, and we must design the service measure so that it is less sensitive to reporting bias.

### **2.3 Algorithmic Fairness in Data Driven Government**

There is a wider risk when cities feed biased reporting data into models that guide where resources go. Lum and Isaac show, in the case of predictive policing, that a model trained on biased records can send resources back to the same places, which produces more records and reinforces the original bias in a feedback loop [10]. The same risk applies to service allocation based on 311 volume. If a city sends crews mainly to neighborhoods that complain the most, it may starve quieter neighborhoods that have real but unreported problems. Our framework is designed in part to expose this risk, by separating measured service response from raw reporting volume and by flagging areas that are quiet but burdened.

### **2.4 Spatial Statistics for Detecting Clusters**

To decide whether disadvantage is concentrated rather than scattered, we use standard tools from spatial analysis. Moran's  $I$  measures whether similar values cluster together across space [11]. Getis and Ord introduce distance based statistics that identify local hot spots and cold spots of high or low values [12]. Anselin generalizes these ideas into local indicators of spatial association, often called LISA, which break a global measure such as Moran's  $I$  into the contribution of each location and reveal local clusters and spatial outliers [13]. These methods

are well established, are available in open source software, and are well suited to tract level data. We use them to test for spatial concentration and to map the specific clusters that define neglected areas.

### **2.5 Open Data Infrastructure and Its Fragility**

Our framework depends on federal environmental data, and recent events show that this data can be fragile. In July 2024 the EPA released EJScreen version 2.3, which added indicators such as nitrogen dioxide levels, drinking water non compliance, extreme heat, and air toxics cancer risk [14]. In early February 2025 the EPA removed public access to EJScreen from its website [15]. A volunteer coalition, the Public Environmental Data Partners, reconstructed version 2.3 and made it publicly available so that research and advocacy could continue [15]. Several archives also preserved the underlying datasets [15]. This episode is more than a footnote. It shows that the data needed to study equity can be withdrawn, and it makes a strong case for building analyses that are reproducible from archived and openly licensed sources, and for documenting data provenance with care.

## **3. Data**

The framework uses four public data sources. All are free, machine readable, and downloadable in bulk. Table 1 summarizes them. We describe each below and note its role and its limits.

311 service requests. City open data portals publish 311 records with a complaint type, a location, a creation timestamp, and usually a closure timestamp. New York City publishes its full 311 history through NYC Open Data, with millions of records each year. From these fields we derive, for each request, the resolution time as the gap between creation and closure, and we place each request inside a census tract using its coordinates. The same structure exists in other cities, such as the Chicago Data Portal, which makes the framework transferable.

American Community Survey. The United States Census Bureau publishes the American Community Survey, or ACS, with five year estimates at the census tract level. From the ACS we take median household income, the racial and ethnic composition of each tract, the share of residents aged sixty five and over, the share of

households with limited English, and population density. These variables describe the social profile of each tract and form the basis of the vulnerability measure.

EJScreen environmental indicators. The EPA EJScreen tool reports a set of environmental burden indicators for small areas across the country, including fine particulate matter, ozone, nitrogen dioxide, traffic proximity, and proximity to hazardous waste facilities, among others [14]. Because the EPA removed direct public access in 2025, we use the reconstructed version 2.3 hosted by the Public Environmental

Data Partners, and we record the access date and the version for reproducibility [15]. EJScreen data are reported at the block group level, so they must be aggregated or crosswalked to the tract level.

Census geography. The Census Bureau publishes TIGER/Line shapefiles with the boundaries of every census tract. We use these geometries to place 311 points into tracts, to crosswalk block group data to tracts, and to build the spatial weights matrix needed for the spatial statistics. The TIGER tract identifier, the GEOID, also serves as the join key to the ACS.

**Table 1. Public Data Sources Used in the Framework. All Sources Are Free and Machine Readable.**

| Dataset                            | Source (free)                                              | Spatial unit                      | Key variables used                                                           |
|------------------------------------|------------------------------------------------------------|-----------------------------------|------------------------------------------------------------------------------|
| 311 service requests               | City open data portal (NYC Open Data, Chicago Data Portal) | Point, joined to tract            | Complaint type, location, created and closed timestamps                      |
| American Community Survey (5 year) | U.S. Census Bureau                                         | Census tract                      | Median income, race and ethnicity, age 65+, limited English, density         |
| EJScreen indicators                | v2.3 EPA, via Public Environmental Data Partners mirror    | Block group, crosswalked to tract | PM2.5, ozone, NO <sub>2</sub> , traffic proximity, hazardous waste proximity |
| TIGER/Line boundaries              | U.S. Census Bureau                                         | Census tract                      | Tract geometries and GEOID for joins and weights                             |

## 4. Methods

### 4.1 Spatial Unit and Data Integration

We use the census tract as the unit of analysis. Tracts are small enough to capture neighborhood level variation and large enough to give stable ACS estimates. The integration proceeds in three steps. First, we assign each 311 request to a tract by a point in polygon join against the TIGER/Line boundaries. Second, we crosswalk EJScreen block group indicators to tracts, using a population weighted aggregation so that larger block groups carry more weight. Third, we merge the ACS variables, which are already at the tract level, by GEOID. The result is a single table with one row per tract. We record the year of each source and the EJScreen version so that the merged table can be rebuilt exactly.

### 4.2 Service, Burden, and Vulnerability Measures

For each tract we compute several service measures from the 311 records. The first is complaint volume, the count of requests. The second is the average resolution time in days. The third is the median resolution time, which is more robust to extreme cases. The fourth is the unresolved rate, the share of requests still open at the time of download. We prefer resolution

time and unresolved rate over raw volume as service signals, because they describe what the city does with a request, while volume mostly describes who chooses to report. The environmental burden measure is a composite. We standardize each EJScreen indicator and average them into a single burden index. The social vulnerability measure is also a composite, built from the ACS variables by standardizing and averaging indicators of low income, minority share, older population, and limited English. Each measure is also reported on its own, so that no single index hides the pattern beneath it.

### 4.3 Testing for Association and Spatial Concentration

We first ask whether burden and weak service are associated with social vulnerability. We begin with simple bivariate relationships and then estimate ordinary least squares regressions of response time and of burden on the demographic variables, with controls for population density and request mix. Ordinary least squares assumes that residuals are independent, which is rarely true for spatial data, since nearby tracts tend to resemble each other. We therefore test the regression residuals for spatial autocorrelation using global Moran's I [11]. If the residuals are spatially autocorrelated,

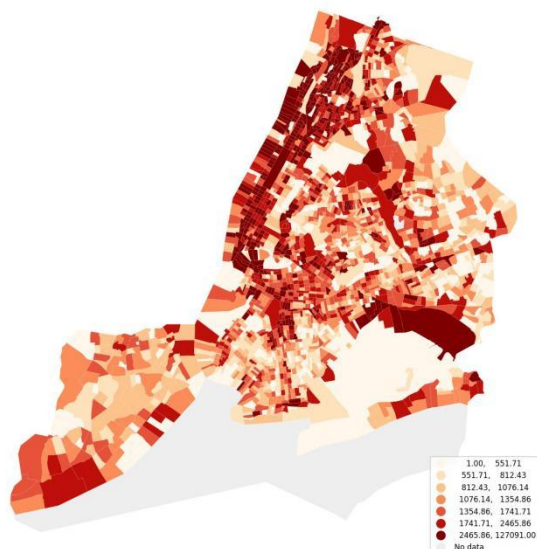
we move to spatial regression models, namely the spatial lag model and the spatial error model, which account for the influence of neighboring tracts. To locate clusters, we compute local Moran's I, or LISA, and the Getis-Ord statistic for each measure, which reveal hot spots and cold spots [12], [13]. A bivariate local Moran's I finds tracts where high burden coincides with high vulnerability in the same place, which is the spatial signature of injustice.

#### 4.4 Identifying Systematically Neglected Tracts

We define a systematically neglected tract through the combination of high environmental burden, weak service response, and high social vulnerability. In the New York City run reported below, we implement the service part of this definition as a neglect score. We turn three service signals, complaint volume, average resolution time, and unresolved rate, into percentile ranks across all tracts. A higher rank means more demand, slower service, or more unresolved cases. We then combine the three ranks with equal weight into one score scaled from zero to one, where a higher score is meant to mark a more neglected tract. We flag the tracts with the highest scores. This first version uses 311 signals only. The environmental burden layer and the demographic layer are not yet joined in this run. As the results show, this matters, and Section 6.4 explains why volume should not be used in this way without correction.

### 5. Application to New York City: Results

Map 1 — Complaint volume by tract (NYC 311, 2025 partial)



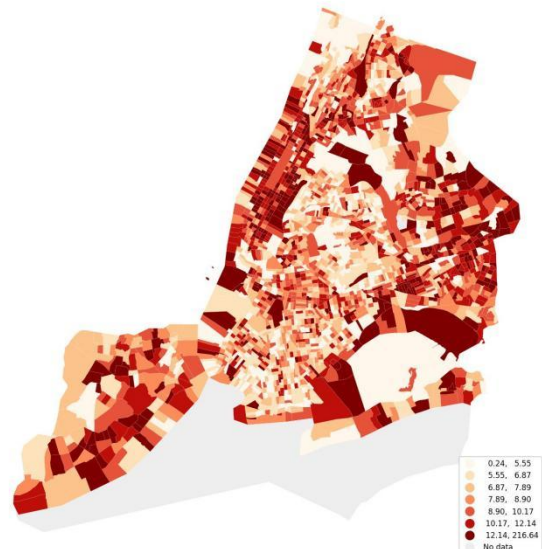
#### 5.1 the Run

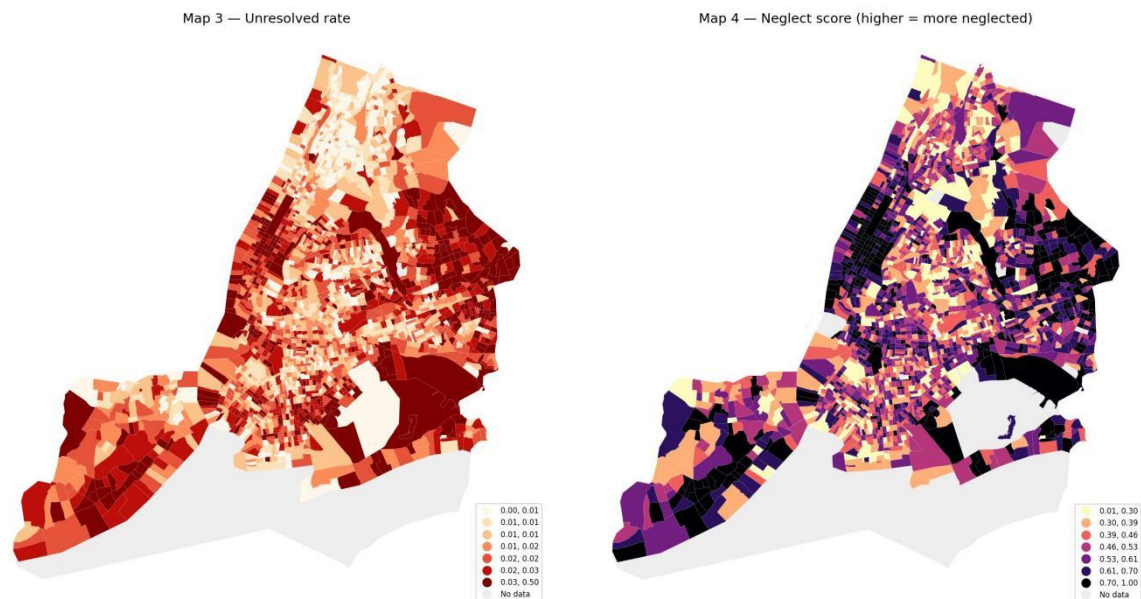
We applied the service part of the framework to New York City using one partial year of 311 service requests from 2025, published through NYC Open Data, joined to 2020 census tracts. The pipeline assigned each request to a tract, computed complaint volume, average and median resolution time, and the unresolved rate per tract, and then built the neglect score described in Section 4.4. We stress again that this run uses 311 service signals only. The EJScreen burden layer and the ACS demographic layer, which the framework calls for, were not joined in this run. Adding them is the next step, and we return to this point in Section 5.4.

#### 5.2 Citywide Patterns

Figure 1 shows the four tract level maps. Map 1 shows complaint volume. Volume is highly uneven across the city. It ranges from a single request in the quietest tracts to several thousand in the busiest central tracts, with one extreme tract above one hundred thousand. Map 2 shows average resolution time, which ranges from under one day in the fastest tracts to well over one hundred days in the slowest, with a small number of tracts far above that. Map 3 shows the unresolved rate, which reaches about one half in the worst tracts. Map 4 shows the combined neglect score on a scale from zero to one. The darkest tracts in Map 4, which carry the highest scores, sit mostly in the dense central areas, with some in the outer boroughs.

Map 2 — Average resolution time (days)





**Figure 1. Tract Level Results for New York City from One Partial Year of 311 Data, 2025. Map 1 (top Left) Shows Complaint Volume. Map 2 (top Right) Shows Average Resolution Time in Days. Map 3 (bottom Left) Shows the Unresolved Rate. Map 4 (bottom Right) Shows the Combined Neglect Score, Where Darker Means More Neglected. Grey Tracts Have No Data.**

One clear feature of the data is that average resolution time is shaped by a small number of very slow cases. Among the most flagged tracts, twenty one of twenty five have an average resolution time more than three times their median. This means that in most of these tracts, half of the requests are closed quickly, while a few requests stay open for a very long time and pull the average up. The median is therefore a more honest measure of the typical experience. The gap between the mean and the median is itself a useful signal that a few cases are stuck.

### 5.3 the Most Neglected Tracts

The pipeline flagged twenty five tracts with the highest neglect scores, with scores from 0.91 to 1.00. Table 2 lists the ten highest. The single highest scoring tract is in Queens, with 5,572 requests, an average resolution time of 179 days, a median of 153 days, and an unresolved rate of 0.42. This tract is a real case of slow and incomplete service, since even the median request takes about five months to close. The tract with the highest unresolved rate is in Manhattan, where almost half of all requests, 0.50, were still open at the time of download.

**Table 2. The Ten Tracts with the Highest Neglect Scores.**

| Tract (GEOID) | Borough   | Complaints | Avg days | Median days | Unresolved | Score |
|---------------|-----------|------------|----------|-------------|------------|-------|
| 36081006201   | Queens    | 5,572      | 179.0    | 152.8       | 0.415      | 0.997 |
| 36061015700   | Manhattan | 6,017      | 81.4     | 3.0         | 0.335      | 0.997 |
| 36081071600   | Queens    | 4,194      | 82.4     | 60.4        | 0.320      | 0.992 |
| 36047020700   | Brooklyn  | 4,073      | 80.4     | 2.9         | 0.309      | 0.991 |
| 36061006600   | Manhattan | 4,889      | 28.4     | 0.1         | 0.136      | 0.990 |
| 36061015301   | Manhattan | 3,321      | 129.5    | 47.4        | 0.498      | 0.987 |
| 36061015302   | Manhattan | 3,248      | 47.2     | 0.4         | 0.277      | 0.983 |
| 36047007500   | Brooklyn  | 2,713      | 58.1     | 0.9         | 0.119      | 0.973 |
| 36061018700   | Manhattan | 2,907      | 19.2     | 1.1         | 0.063      | 0.960 |
| 36081005800   | Queens    | 2,184      | 40.1     | 0.5         | 0.106      | 0.957 |

The Full Set of Twenty Five Flagged Tracts Is Provided in the Accompanying File Neglected\_tracts.csv. Avg Days Is the Average Resolution Time, Median Days Is the Median Resolution Time, and Unresolved Is the Share of Requests Still Open.

The borough breakdown of the twenty five flagged tracts is striking. Seventeen are in Manhattan, six are in Queens, and two are in Brooklyn. None are in the Bronx, and none are in Staten Island. This result needs careful reading, which we give next.

#### 5.4 Reading the Result with Care

At first sight, a list of neglected tracts that is mostly in Manhattan is hard to believe. Manhattan is the wealthiest borough and is generally well served. The reason for the pattern is built into the score. The neglect score uses complaint volume as one of its three parts, and complaint volume is highest in dense central areas where many residents file many requests. As the literature on 311 reporting shows, the number of complaints reflects the propensity to report, which rises with density, income, civic engagement, and English language ability, and not only the true level of unmet need [5], [9]. So the volume part of the score pulls the result toward Manhattan, where reporting is heaviest.

The maps make this clear. Map 2 and Map 3 show that slow resolution and high unresolved rates also appear in outer areas, including parts of the eastern Bronx, southeast Queens, and Staten Island. These are areas that prior environmental justice work would expect to be under served. But because these areas file fewer complaints, the volume weighted score does not rank them at the top. The result is therefore a clear and live example of the reporting bias that this paper warned about. It shows in practice that raw complaint volume should not be treated as a sign of neglect.

The correction follows directly from the framework. First, the volume signal should be dropped or down weighted, or normalized by population or by an estimate of true need, so that quiet but burdened tracts are not hidden. Second, the environmental burden layer from EJScreen and the demographic layer from the ACS should be joined, so that the analysis can separate true need from the propensity to report. With those layers in place, a high neglect score would mean high burden and weak service in a vulnerable population, rather than simply a high volume of complaints. This is the planned next step. The present run is best read as the service responsiveness pillar of the full framework, together with a useful warning about how not to combine the signals.

#### 6. Discussion

The framework has direct uses for policy. A city agency could use the flagged list to target investment, to schedule proactive inspections in quiet but burdened areas, and to build an accountability dashboard that tracks whether

gaps narrow over time. Because the method is transparent, residents and oversight bodies can check the result and challenge it, which supports the accountability goal of open data. The New York City run also shows the value of a careful design. The slow and unresolved tracts in Table 2, especially the ones where the median request takes months to close, are concrete cases that an agency could act on at once.

The case study also gives a sharp warning, which the related work predicts and which our own numbers confirm. A city should not allocate resources on the basis of raw 311 volume, because doing so rewards the neighborhoods that already complain the most and can starve quieter neighborhoods with real needs. This is the same feedback loop that Lum and Isaac describe for predictive policing [10]. Our flagged list, with seventeen of twenty five tracts in Manhattan, is what this loop looks like in practice. The fix is to lead with service measures such as resolution time and unresolved rate, which describe the city's own action, and to bring in burden and demographic data so that need can be told apart from the propensity to report [5], [9].

A third lesson concerns the data itself. The removal of EJScreen from the EPA website in 2025 shows that equity analysis can lose its data sources for reasons that have nothing to do with the science [15]. This is a practical risk for any researcher who builds on federal environmental data, and it is also a point about justice. When the public loses access to data that documents unequal burdens, the communities that bear those burdens lose a tool for advocacy. A reproducible pipeline that records its sources, that can run from archived and openly licensed copies, and that documents version and access dates, is therefore not only good engineering. It is part of keeping the analysis possible at all. We recommend that studies in this area cite the specific mirror or archive used and keep a local copy of the input data.

Finally, the framework is portable. The same pipeline runs on any city that publishes 311 records and that lies within EJScreen and ACS coverage. Comparing several cities would show whether the patterns are general or local, and would let researchers ask why some cities show smaller gaps than others.

#### 7. Limitations

Several limitations should be kept in mind. First, the New York City run uses 311 service signals

only. The EJScreen burden layer and the ACS demographic layer are not yet joined, so the present neglect score measures service side neglect, not full environmental justice. Second, the score includes raw complaint volume, which is sensitive to reporting bias, and the flagged list reflects that bias rather than true need. We report this openly because it is the central lesson of the run. Third, average resolution time is skewed by a small number of very slow cases, so we report the median alongside it. Fourth, the data is one partial year of 2025, and the unresolved rate is affected by the download cutoff, because recent requests have had less time to close. Fifth, the analysis is at the tract level, so it describes places and not individuals, which is the ecological inference problem, and results can depend on the choice of spatial unit, which is the modifiable areal unit problem [2]. Sixth, the design is cross sectional, so it describes association and concentration rather than cause.

## **8. Conclusion**

Open government data makes it possible to ask, in a concrete and testable way, whether a city treats its neighborhoods fairly. This paper has set out a reproducible framework that joins 311 service requests, Census demographics, and EPA environmental indicators at the tract level, and that uses established spatial statistics to find neighborhoods that are systematically neglected. We applied the service part of the framework to New York City and produced tract level maps of complaint volume, resolution time, unresolved rate, and a combined neglect score, together with a list of the twenty five most neglected tracts. The result carries a clear methodological message. Because the score included raw complaint volume, it flagged mostly wealthy central tracts, which is a live example of reporting bias rather than a finding about true neglect. The cure is to down weight volume and to add the environmental and demographic layers, so that need can be separated from the propensity to report. The pipeline is ready for that next step, and it can be run on any city with a 311 system. The episode of EJScreen, removed by the EPA and rebuilt by volunteers in 2025, is a reminder that data preservation is part of equity work.

## **References**

[1]Mohai, P., Pellow, D., and Roberts, J. T. (2009). Environmental Justice. Annual

Review of Environment and Resources, 34, 405-430.

[2]Mohai, P., and Saha, R. (2006). Reassessing Racial and Socioeconomic Disparities in Environmental Justice Research. *Demography*, 43(2), 383-399.

[3]Morello-Frosch, R., Zuk, M., Jerrett, M., Shamasunder, B., and Kyle, A. D. (2011). Understanding the Cumulative Impacts of Inequalities in Environmental Health: Implications for Policy. *Health Affairs*, 30(5), 879-887.

[4]Wolch, J. R., Byrne, J., and Newell, J. P. (2014). Urban Green Space, Public Health, and Environmental Justice: The Challenge of Making Cities Just Green Enough. *Landscape and Urban Planning*, 125, 234-244.

[5]Kontokosta, C., Hong, B., and Korsberg, K. (2017). Equity in 311 Reporting: Understanding Socio-Spatial Differentials in the Propensity to Complain. Bloomberg Data for Good Exchange Conference. arXiv:1710.02452.

[6]O'Brien, D. T. (2016). 311 Hotlines, Territoriality, and the Collaborative Maintenance of the Urban Commons. *Evolutionary Behavioral Sciences*, 10(2), 123-141.

[7]Wang, L., Qian, C., Kats, P., Kontokosta, C., and Sobolevsky, S. (2017). Structure of 311 Service Requests as a Signature of Urban Location. *PLOS ONE*, 12(10), e0186314.

[8]Minkoff, S. L. (2016). NYC 311: A Tract-Level Analysis of Citizen-Government Contacting in New York City. *Urban Affairs Review*, 52(2), 211-246.

[9]Liu, Z., Bhandaram, U., and Garg, N. (2024). Quantifying Spatial Under-reporting Disparities in Resident Crowdsourcing. *Nature Computational Science*, 4(1), 57-65.

[10]Lum, K., and Isaac, W. (2016). To Predict and Serve? Significance, 13(5), 14-19.

[11]Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1-2), 17-23.

[12]Getis, A., and Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24(3), 189-206.

[13]Anselin, L. (1995). Local Indicators of Spatial Association - LISA. *Geographical Analysis*, 27(2), 93-115.

[14]U.S. Environmental Protection Agency

(2024). EJScreen: Environmental Justice Screening and Mapping Tool, Technical Documentation for Version 2.3.

[15]Environmental Data and Governance Initiative and Public Environmental Data

Partners (2025). EPA Removes EJScreen From Its Website; Reconstruction of EJScreen Version 2.3. Public Environmental Data Partners, [screening-tools.com](https://screening-tools.com).