Neighborhood Neglect: Examining Inequity and Inequality of Boston 311 Service Requests

Ethan Long

Table of contents

1 Abstract

In this study, we aim to examine the underlying patterns and statistical drivers behind resolution times for Boston 311 Service Requests. Boston 311 serves as a non emergency hotline for any resident who wishes to report neighborhood issues via phone or by mobile app. Since its adoption in 2011, it has become a vital data source for urban research and city development. Many major U.S. cities including New York City, Chicago, Philadelphia, and Denver have since implemented similar systems.

Using the publicly available Boston 311 data spanning 2011 to present day, we analyze the historical performance of public service in resolving requests to identify possible patterns of inefficiency and inequity. By supplementing this dataset with neighborhood-level socioeconomic context from the American Community Survey (ACS), this study aims to pinpoint areas of the city that may be under served or subject to neglect.

2 Introduction and Related Literature

This study seeks to determine whether the 311 system can reveal broader patterns of residential issues in Boston. We will also analyze temporal and spatial trends in request submissions and resolution times, identify demographic and socioeconomic factors associated with service disparities, and provide insights for improving municipal responsiveness and equity.

Additionally, recent statistical work has demonstrated that 311 service request data contains rich spatial-temporal structures suitable for modeling municipal performance and urban inequities (Yan et al., 2020).

This study builds on a growing body of research that uses 311 service request data as a proxy for neighborhood conditions and urban inequalities. Prior work done by Wang et al. (2017) demonstrated that the categorical structure of 311 requests can reveal distinctive socioeconomic signatures across neighborhoods, motivating the integration of service request data with census-level indicators.

More broadly, urban scholars have emphasized the role of municipal service delivery as a mechanism, through which spatial inequality is reproduced (Batty, 2013; Kontokosta, 2015). Differences in response times and service quality have been shown to correlate with income, racial composition, and political influence, raising concerns about equity in urban governance.

From a methodological perspective, prior studies caution against purely predictive approaches when modeling complex urban systems characterized by heavy-tailed distributions and spatial dependence (Bettencourt, 2013). Consequently, recent research advocates combining machine learning methods with interpretable statistical frameworks to better understand structural drivers of observed disparities.

This paper contributes to this literature by examining resolution time disparities within Boston's 311 system using a hybrid analytical approach that emphasizes both predictive performance and interoperability.

3 Data and Preprocessing

3.1 Data Wrangling

Boston 311 data is a publicly available data source that can be accessed through either manual queries or via the Boston 311 API. For this study, annual service request data was acquired through manual queries, as this approach was a more efficient and consistent method of data collection.

We combined the existing features of the 311 dataset with a series of engineered temporal and spatial features to enhance analytical depth. these were then integrated with demographic and socioeconomic context from the American Community Survey (ACS) to develop a comprehensive view of how certain neighborhoods are served. This fusion of service and demographic data allows for a systematic examination of potential disparities in municipal responsiveness across Boston's communities.

Prior research has shown that integrating 311 service logs with demographic and spatial features gives reliable statistical insight into neighborhood level disparities (Yan et al., 2020).

Boston's public service sector has been the subject of controversy over the last hundred years due to past corruption, inequity, and rapidly developing neighborhoods. These factors lead to systemic issues becoming rooted in all facets of life and beg to be examined.

Boston's public service landscape provides a unique context for this analysis. Historically shaped by waves of rapid urban development, demographic change, and periods of political corruption, the city's service delivery patterns may reflect long-standing inequities. This study uses 311 data as a lens to quantify and better understand these disparities in modern Boston.

3.2 Data Cleaning

Raw Boston 311 service request records require substantial preprocessing prior to statistical analysis due to missing timestamps, inconsistent geocoding, duplicate records, and extreme variability in service resolution times. All data preparation steps were conducted to ensure reproducibility and to minimize bias introduced by data sparsity or reporting artifacts.

Service requests were first filtered to retain only observations with valid opening and closing timestamps. Resolution time was computed in hours as the difference between the service request closure and opening times. Requests with non-positive resolution durations were removed, as these likely reflect data entry errors or automated closures rather than meaningful service activity.

Given the highly right-skewed distribution of resolution times, extreme outliers were examined carefully. requests exceeding the 99th percentile of resolution duration were retained for descriptive analysis but flagged during model evaluation to assess sensitivity to unusually long service delays. This decision reflects the operational reality that extreme delays, while rare, may disproportionately affect certain neighborhoods and therefore carry substantial importance for equity analysis.

Spatial preprocessing involved mapping individual service requests to census tracts using geographic coordinates. Census tract identifiers were then used to merger neighborhood-level demographic characteristics from the American Community Survey (ACS), including median household incomes, educational attainment, and population density. Requests with missing or invalid spatial identifiers were excluded from tract-level analyses.

Temporal features were derived from request timestamps, including hour of the day, day of the week, and month of submission, allowing for us to include the separation of systematic temporal effects from neighborhood level disparities. All preprocessing steps were performed using reproducible scripts to ensure consistency across analyses.

3.2.1 Handling Missing Data

The raw Boston 311 datasets comprise all recorded service requests along with their associated descriptors, including timestamps, geographic coordinates, request type, and responding agency. Because the system processes hundreds of requests daily, many of which are routine or low stakes, several fields were frequently incomplete or inconsistently entered. Commonly missing fields included the 'before/after' photos, the request's target resolution time, and occasionally the reported ZIP code. These variables either contained excessive missingness or were not analytically relevant for modeling municipal response performance. Therefore, they were confidently removed during preprocessing.

Location accuracy, however, is essential for spatial analysis. When ZIP codes were missing, we used reverse geocoding (via geopy) to infer ZIP codes from the available latitude and longitude coordinates. Requests lacking valid geographic coordinates were removed, as they could not be joined to census tracts or ACS data.

Timestamps missing either opn_dt or closed_dt were also excluded, since they prevented the calculation of the primary outcome variable: resolution time (in hours).

3.2.2 Feature Standardization

Several key features required standardization to ensure interpretability and computational overhead during modeling. All timestamp fields were parsed into datetime objects with a consistent timezone, enabling extraction of temporal predictors such as hour of the day, day of week, month, season, and indicators of working-hour activity. Categorical text fields (request type, department) were normalized by trimming whitespace, converting to lowercase, and resolving minor spelling inconsistencies

Spatial fields were standardized by converting latitude and longitude to numeric types and validating their ranges. Each request was then spatially joined to a census tract and ZIP code, facilitating merges with ACS socioeconomic data and enabling computation of geospatial features such as distance to City Hall.

Continuous numeric fields (resolution time) were reviewed for implausible or negative values and corrected or removed when necessary. Extremely large resolution times, often resulting from unresolved cases, were capped and filtered using a high-percentile rule to prevent any distortion of model training.

4 Feature Engineering

Feature engineering efforts were split into two spatial and temporal components to enrich the raw Boston 311 dataset beyond timestamps and coordinates.

4.1 Spatial Features

Spatial transformations aimed to capture geographic context and socioeconomic conditions influencing request handling.

While American Community Survey (ACS) variables were merged at the neighborhood level to contextualize service request patterns, the primary regression models do not include these variables explicitly. Instead, neighborhood fixed effects are used to proxy for underlying socioeconomic and demographic conditions, capturing income, education, housing stock, and historical patterns of municipal investment.

Feature	Type	Description / Purpose
`latitude`, `longitude`	Numeric	Raw geographic coordinates of each service request.
`distance_ch`	Numeric	Vectorized Haversine distance from Boston City Hall, a proxy for spatial accessibility to central services.
`zipcode`, `census_tract`	Categorical	Geographic identifiers used to merge ACS data.
`acs_median_income`, `acs_education`, `acs_population_density`	Numeric	Socioeconomic and demographic indicators from the American Community Survey, joined by ZIP or census tract.

4.2 Temporal Features

Temporal indicators describe the timing of service demand and administrative responsiveness.

Feature	Type	Description / Purpose
`hour`, `day_of_week`, `month`	Numeric	Extracted from timestamps to reveal periodic service patterns.

Feature	Type	Description / Purpose
`working_hours`, `rush_hour`, `weekend`, `season`	Binary / Categorical	Flags capturing business-hour and seasonal effects on resolution time.
`time_elapsed_hours_log`	Numeric	Resolution time (target variable), computed as the hour difference between `closed_dt` and `open_dt`.

5 Statistical Modeling Framework

To estimate the association between neighborhood characteristics and service resolution times, we first fit a linear regression model to the log-transformed outcome. Predictor variables include distance to Boston's City Hall, temporal indicators, department-level fixed effects, and neighborhood fixed effects. Neighborhood fixed proxy for underlying socioeconomic and demographic conditions, allowing the model to capture spatial inequities without explicitly including census covariates.

This modeling framework enables direct statistical inference on the direction and magnitude of associations between neighborhood characteristics and expected service delays, complementing the predictive machine learning models presented later in the analysis.

5.1 Outcome Definition and Transformation

5.1.1 Distributional Properties of Resolution Time

Resolution times for Boston 311 service requests exhibit substantial right skew and heavy-tailed behavior, with a small fraction of requests remaining open for extended durations. Such distributions violate the constant-variance and symmetry assumptions underlying standard linear regression when modeled on a raw scale.

To address this issue, resolution time was log-transformed prior to statistical modeling. Specifically, the transformed outcome was defined as

$$Y_i = \log(T_i + 1), \tag{1}$$

where T_i denotes the resolution time in hours for request i. The addition of one unit ensures numerical stability for rapidly resolved requests while preserving relative differences among longer delays. All inferential models described below were estimated using this transformed outcome.

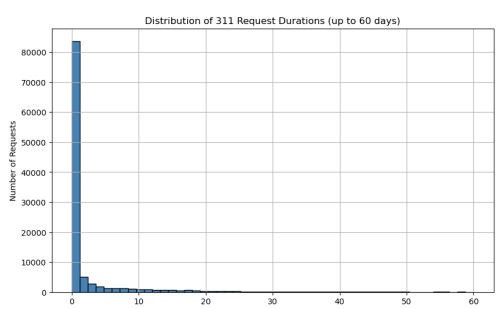

Figure 1: Distribution of Boston 311 service request resolution times on the raw scale. The raw distribution exhibits substantial right skew and heavy-tailed behavior, motivating the use of a logarithmic transformation for statistical modeling.

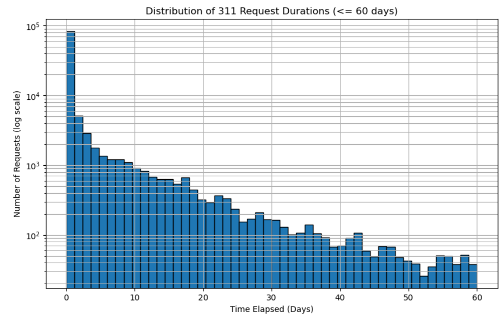

Figure 2: Distribution of log-transformed Boston 311 service request resolution times.

5.2 Linear Regression Model

To assess the association between spatial location, neighborhood context, and service resolution times, we estimate a linear regression model with the log-transformed resolution duration as the outcome variable. The model includes controls for distance to Boston's City Hall, temporal indicators (weekend, rush hour, working hours), department-level fixed objects, and neighborhood fixed effects. Robust Standard errors (HC3) are used to account for heteroskedasticity.

Neighborhood fixed effects are included to proxy for underlying socioeconomic and demographic differences across areas of the city, capturing variation in income, education, housing stock, and historical patterns of municipal investment that are not directly observed in the service request data.

6 Results

6.1 Linear Regression Results

Table reports the estimated coefficients from the linear regression model. The model explains approximately 13.6% of the variation in log-transformed resolution times (R^2 = 0.136), a magnitude consistent with prior studies of administrative service systems characterized by substantial case-level heterogeneity.

Spatial location is a statistically significant predictor of service delay. Distance from Boston City Hall is positively associated with longer resolution times ($\hat{\beta} = 2.74 \times 10^{-5}, p < 0.001$), indicating that requests submitted farther from the city center experience systematically longer delays even after accounting for department assignment, neighborhood, and timing.

Temporal patterns further reveal operational constraints. Requests submitted during rush hours and standard working hours exhibit significantly longer expected resolution times, while weekend submissions are resolved more quickly on average, likely due to automated closures or differential prioritization of requests.

Substantial heterogeneity is observed across municipal departments. Agencies involved in infrastructure maintenance, inspections, and housing services exhibit markedly longer resolution times relative to the baseline department, underscoring the role of institutional processes in shaping service efficiency.

Neighborhood fixed effects reveal persistent spatial disparities. Central and higher-income neighborhoods such as Back Bay, Beacon Hill, the South End, and South Boston experience significantly shorter resolution times, while neighborhoods including Mattapan, Roxbury, and Dorchester exhibit longer delays, even after controlling for departmental and temporal factors. These patterns suggest that operational inequities in municipal service delivery align with broader spatial and socioeconomic divides within Boston. Because neighborhood fixed effects absorb socioeconomic variation, these differences should be interpreted as structural spatial disparities rather than causal effects of any single demographic characteristic.

6.2 Predictive Performance Benchmarks

In addition to the inferential regression framework, Random Forest and Gradient Boosting models were estimated to assess the extent to which nonlinear interactions and complex feature relationships could improve predictive performance. These models are not intended for real-time forecasting, but rather to evaluate whether service delays can be reliably predicted from observable neighborhood and temporal features alone.

The resulting predictive performance was modest ($R^2 = 0.214$ for Random Forests and $R^2 = 0.182$ for Gradient Boosting), suggesting that a substantial portion of variation in resolution times arises from unobserved operational factors, discretionary decision making, or idiosyncratic case level dynamics.

The relatively modest R^2 reflects the inherently random and case-specific nature of municipal service requests, which depend on unobserved factors such as staffing availability, request prioritization, and inter-agency coordination.

6.3 Feature Importance and SHAP Analysis

Feature importance and SHAP values are derived from tree-based models estimated for exploratory purposes. In these models, neighborhood indicators and spatial variables implicitly encode socioeconomic variation, allowing assessment of how location-based characteristics contribute to predictive performance.

1. Distance from Boston City Hall (km)
2. Responsible Department/Agency
3. Request Time (Service Category)
4. Neighborhood Indicators
5. Temporal Features (Hour/Day)

Departmental assignment dominated model variance (>35% of total importance), reflecting structural differences in case-management processes among divisions such as Public Works, Transportation, and Inspectional Services.

Service category ranked second, confirming that certain complaint types, particularly those involved construction, utilities, or road maintenance, carry inherently longer turnaround times.

Spatially the 'distance_ch' variable contributed nonlinearly, where requests farther from Boston's City Hall correlated with modest but systematic increases in expected delay.

6.4 Spatial Correlation Results

Global Moran's I = 0.247 ($p < 0.0001$) indicated significant positive spatial autocorrelation in resolution times.

Hot-spot analysis showed prolonged delays concentrated in Dorchester, Mattapan, Roxbury, and Hyde Park, while Back Bay, Beacon Hill, Downtown, and South Boston exhibited persistently faster closures.

These clusters of neighborhoods align with lower median-income tracts and higher population densities, supporting the hypothesis that operational inequities mirror socioeconomic and spatial divides.

6.5 Exploratory Spatial and Residual Analysis

A spatial heat map of all Boston 311 service requests was generated to visualize the geographic distribution of requests across the city. The heat intensity indicates the relative frequency of reports within each neighborhood. High density areas correspond to regions with frequent complaints and requests, often near central and high-population neighborhoods, while cool colored areas indicate less service activity.

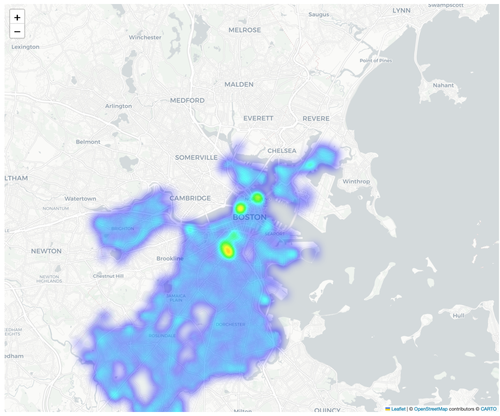

Figure 3: Spatial distribution of average model residuals by neighborhood. Positive values indicate longer-than expected resolution times after controlling for department, timing, and location.

The map below provides a visual baseline of reporting intensity across Boston neighborhoods, serving as a foundation for later analysis of whether these spatial patterns correspond to differences in resolution times or socioeconomic variables.

7 Discussion

7.1 Structural vs Predictable Inefficiency

While machine learning models achieved modest predictive performance, the inferential regression results indicate that service delays are not random. Instead, resolution times are shaped by structural factors including department assignment, spatial location, and neighborhood context. The limited predictability underscores the role of unobserved operational constraints rather than model inadequacy. The presence of neighborhood and department fixed effects increases the condition number of the design matrix but does not bias coefficient estimates, and robust standard errors mitigate concerns related to heteroskedasticity.

7.2 Institutional Sources of Delay

Department-level fixed effects explain a substantial share of systematic variation in resolution times. Agencies responsible for inspections, housing, and infrastructure maintenance exhibit significantly longer delays, reflecting the procedural complexity and coordinating demands inherent to these services. These institutional differences suggest that improving municipal responsiveness requires operational reforms tailored to specific agencies rather than uniform policy interventions.

7.3 Spatial Inequity and Service Equity

Even after controlling for department and temporal factors, neighborhood-level disparities persist. Requests originating from neighborhoods such as Mattapan, Roxbury, and Dorchester experience longer delays relative to central and higher-income areas. Because neighborhood fixed effects proxy for underlying socioeconomic and historical conditions, these findings suggest that municipal service delivery may reinforce existing spatial inequities. These disparities persist even when socioeconomic conditions are modeled indirectly via neighborhood fixed effects, suggesting that observed inequities reflect systemic differences in municipal responsiveness rather than isolated demographic factors.

8 Conclusions

This study examines inequities in Boston's 311 service request system using a hybrid statistical and machine learning framework. By modeling service resolution times as a function of spatial, temporal, institutional, and neighborhood-level factors, we identify persistent disparities in municipal responsiveness.

The results demonstrate that service delays are not uniformly distributed across the city. Instead, they reflect institutional constraints and spatial patterns that align with broader socioeconomic divides. While predictive accuracy is inherently limited by unobserved operational factors, inferential modeling reveals systematic inequities that warrant policy attention.

More broadly, this analysis illustrates how administrative service data can be used not only to monitor municipal performance but also to diagnose structural inequalities in urban governance. Future work could integrate staffing data, service-level agreements, or budget allocations to further disentangle the mechanisms underlying observed disparities.

9 Bibliography

Batty, M. (2013). The new science of cities. MIT Press.

Bettencourt, L. M. A. (2013). The origins of scaling in cities. Science, 340(6139), 1438–1441. https://doi.org/10.1126/science.1235823

Kontokosta, C. E. (2015). The quantified community and neighborhood labs: A framework for computational urban science and civic technology innovation. Journal of Urban Technology, 22(3), 67–84. https://doi.org/10.1080/10630732.2015.1072182

Wang, Y., Chen, J., & Xie, Y. (2017). Classifying urban service requests using semantic information. ACM Transactions on Intelligent Systems and Technology, 8(4), 1–19. https://doi.org/10.1145/3041023

Yan, J., Chen, J., & Xie, Y. (2020). A statistical framework for analyzing city service request data. *Journal of the American Statistical Association*, *115*(532), 1889–1904. https://doi.org/10.1080/01621459.2020.1765783