

## **Enhancing NYC 311:**

### **Identifying Key Drivers of Dissatisfaction and Trend**

Group 7: Minkyung Kim, Boni Vasiu Rosen, Minqi Huang, Jiarong Ying

APANS5205

Aug 11, 2025

# Introduction

In modern urban governance, effective communication between citizens and public service agencies plays an essential role in maintaining trust and enhancing the quality of life across the city. In New York City, the 311 system functions as the central non-emergency reporting platform, which enables residents to voice their concerns, ranging from noise complaints and illegal parking to sanitation and infrastructure problems. As one of the largest and most active civic engagement systems in the world, NYC311 serves as both a bridge of public satisfaction and municipal responsiveness.

However, according to recent studies, the latest trends point to a growing mismatch between resident expectations and agency outcomes. As complaint volumes increase while satisfaction scores remain consistently low, questions have emerged about the effectiveness of city agencies' responses – particularly in the way service resolutions are communicated. Under these circumstances, while operational data, such as complaint types and request times, offer valuable metrics, less attention has been paid to the textual narratives that accompany resolution outcomes. This field of narratives, which is often drafted by agency personnel, may play a crucial yet understudied role in shaping citizen perceptions of service quality and agency accountability.

This project takes a dual approach to explore dissatisfaction in NYC311 services, including the relationship between the language used in NYC311 resolution descriptions and reported citizen satisfaction. First, we analyze temporal trends in complaint volumes to understand seasonal surges and shifting public concerns, using time series forecasting methods to highlight recurring patterns and anticipate future demands. Second, we investigate the relationship between resolution description language and citizen satisfaction, evaluating whether certain linguistic patterns, such as vague phrasing or boilerplate language, correlate with dissatisfaction. In this way, by integrating quantitative complaint data with qualitative text analysis, we aim to generate actionable insights for improving the transparency, efficiency, and accountability of NYC311 services.

Our analysis relies on two large-scale datasets from NYC Open Data: the “311 Service Requests from 2022 to Present” and the “311 Resolution Satisfaction Survey.” By merging these data sources and applying various data analysis techniques, we aim to better understand the dynamics of dissatisfaction, uncover patterns of ineffective communication, and recommend strategies for enhancing public service outcomes across different boroughs and complaint types.

## Literature Review

Understanding the drivers of public dissatisfaction in large-scale civic service systems requires a multidisciplinary approach—drawing on research in urban analytics, behavioral public administration, text analysis, and temporal trend modeling. Prior studies on 311 systems and citizen engagement have highlighted the importance of timely responses and operational follow-ups. However, emerging evidence suggests that the *language* used in service communications may be just as critical in shaping public trust and perceived effectiveness. This

section synthesizes key findings from recent studies relevant to our research focus, particularly how complaint frequency, agency communication styles, and systemic service disparities contribute to patterns of dissatisfaction in the NYC311 system.

### *The Rise of NYC311 Complaints and Their Geospatial Clustering*

Recent data highlights a steady increase in non-emergency complaints submitted to NYC311, which reflects growing public concerns about New York City's living conditions and service quality. According to DiNapoli's (2025) findings, NYC311 received over 3.4 million calls in 2024, representing a 7% increase from 2023, with notable surges in complaints about illegal parking (+155% since 2019), noise disturbances (+19% from 2023), and heating or hot water issues (+14% since 2019). These complaints are not evenly distributed; they tend to cluster in high-density or underserved areas such as the Bronx and Downtown Brooklyn. In addition, the NYC311 call data show a highly skewed distribution of complaint types – the top 20 issues account for approximately 70% of all requests (Hussey & Yan, 2025). This suggests that public dissatisfaction is both a temporal and spatial phenomenon, prompting the need for analysis methods that can detect both structural inequities and evolving patterns of concern.

### *Low Satisfaction and the Role of Vague Resolutions*

Despite the scale of NYC311, overall user satisfaction remains remarkably low. As of mid-2025, the NYC311 Resolution Satisfaction Dashboard reported an average score of just 36 out of 100, with a majority of users expressing strong dissatisfaction with their case outcomes (NYC311, 2025). Several studies link this dissatisfaction not solely to the outcomes themselves, but to how those outcomes are conveyed. For instance, W42ST (2025) found that agencies with vague or discretionary resolution language – such as the NYPD's use of noncommittal phrases in illegal parking cases – received some of the lowest satisfaction scores. In contrast, resolutions that included explicit action, like the issuance of a summons, correlated with 80% higher satisfaction, even when the objective outcome was similar. These insights emphasize the need to evaluate the linguistic framing of agency responses as a key determinant of public perception – an area this project addresses using topic modeling and other text analysis techniques.

### *Cognitive Models of Service Dissatisfaction*

From a psychological perspective, dissatisfaction with public services often reflects deeper cognitive and emotional responses to perceived inaction or ambiguity. Research by Lee et al. (2025) introduces an expectation-oriented model which posits that negative experiences – particularly those involving unclear communication – carry more weight than positive ones. This aligns with broader findings in behavioral economics and service quality literature: citizens tend to remember and react more strongly to service failures, especially when those failures appear avoidable or poorly explained.

### *Temporal Analysis and Forecasting in Public Services*

Beyond communication styles, understanding how public service needs shift over time is equally important. Time series analysis has been widely applied in urban systems to anticipate demand

and thus, optimize resource allocation. Research in civic analytics (Liu et al., 2022) indicates the importance of identifying seasonal peaks, such as winter surges in heating-related complaints or summer spikes in noise violations, to help city agencies plan ahead and allocate resources more efficiently. Applying these techniques to NYC311 complaints data allows us to go beyond analysis and develop predictive insights. By decomposing complaint trends into seasonal and trend components, city officials can anticipate public concerns before they escalate and tailor interventions to specific boroughs or complaint types. This complements the linguistic side of our study by adding a strategic and forward-looking layer to our analysis.

The existing literature underscores the critical but frequently overlooked role of resolution narratives in shaping public opinion. The literature suggests that both the content and timing of public service interactions matter. Citizens trust not just on what city agencies do, but also on how and when they respond. Our project contributes to this growing field by integrating natural language analysis with time series modeling to identify pain points in both communication and system responsiveness. We believe that understanding how linguistic cues reflect accountability can help inform smarter and more responsive strategies for improving urban service delivery. We aim to provide a more comprehensive understanding of citizen dissatisfaction and actionable insights for making NYC311 a more transparent and citizen-centered system.

## Research Questions

**RQ1.** What complaint types and borough-level characteristics are associated with low resolution satisfaction or unresolved cases in the 311 system?

- **H1.** Complaints related to infrastructure and law enforcement, such as noise, illegal parking, and sanitation, are more frequent and tend to receive lower satisfaction ratings. These patterns may vary by borough, and Bronx and Brooklyn are expected to show higher dissatisfaction and resolution failure rates.

**RQ2.** How are resolution description language patterns associated with citizen satisfaction, and do these patterns differ by complaint type or responding agency?

- **H1.** Resolution descriptions that include vague, repetitive, or boilerplate phrases such as “no violation observed” and “attempted to inspect” are associated with lower satisfaction ratings.

**RQ3.** How have complaint volumes for different issues changed over time, and what trends or seasonal patterns can be identified from the 311 request data?

- **H1.** Certain complaint types, such as noise, heat/hot water, and sanitation, show consistent seasonal spikes. These trends can inform predictive models that allow agencies to prepare resources more effectively for expected demand.

# Data Description and Suitability

This project draws on two publicly available datasets from the NYC Open Data portal to investigate patterns of dissatisfaction in the NYC311 system: (1) 311 Service Requests from 2022 to Present and (2) 311 Resolution Satisfaction Survey. These datasets allow us to examine complaint types with time and volume, agency responses, linguistic features of resolutions, and satisfaction outcomes across service areas. Together, they offer both structured metadata and unstructured narrative content, which enables a comprehensive analysis.

Because the datasets contain a wide range of information, we narrowed our focus to a smaller set of fields that are most useful for our analysis. These include timestamps, categorical fields such as complaint type, agency, and borough, and free-text resolution descriptions, which enable us to conduct temporal trend analysis, apply text mining techniques, and examine spatial and agency-level patterns in service delivery.

## *311 Service Requests Dataset*

The 311 Service Request dataset captures non-emergency public service complaints submitted by New York City residents. Each entry represents a unique case and includes fields such as: Complaint type and descriptor (e.g., Noise – Loud Music/Party), Agency name and status (e.g., NYPD, Closed), Geographic identifiers (e.g., borough, city), Timestamps for when the complaint was created and closed, and Resolution description, a free-text field written by the agency that summarizes the action taken or the reason for closure. To support trend analysis and temporal modeling, date-time fields were split into separate date and time variables. The dataset was filtered to focus on records from 2022 to the present, aligning with the satisfaction survey's available time frame.

## *311 Resolution Satisfaction Survey Dataset*

The 311 Resolution Satisfaction Survey dataset contains follow-up survey responses in which residents rate their satisfaction with the resolution of their 311 requests. Fields include: Complaint type and borough, Satisfaction response (e.g., Satisfied, Dissatisfied), Reason for dissatisfaction (if provided, in free-text form), and Survey year and month, aggregated to maintain respondent anonymity. Additional preprocessing included cleaning blank strings, normalizing agency and complaint labels, and aligning naming conventions across datasets (e.g., “Resolution.Description” renamed to “Resolution\_Description”).

## *Selected Variables*

After reviewing a wide range of available fields, we narrowed our analysis to a refined subset that best supports the research questions. These include both categorical metadata (e.g., borough, agency, complaint type) and textual features (e.g., resolution descriptions). A summary of selected variables is provided below:

Dataset	Key Variables Used
311 Service Requests	<i>Agency_Name, Complaint_Type, Descriptor, City, Status, Resolution_Description, Borough, Created_Date, Created_Time, Closed_Date, Closed_Time</i>
311 Resolution Satisfaction Survey	<i>Unique_ID, Agency_Acronym, Agency_Name, Complaint_Type, Descriptor, Borough, Resolution_Description, Survey_Year, Survey_Month, Satisfaction_Response, Dissatisfaction_Reason</i>

### *Suitability for Analysis*

The NYC311 datasets are well-suited to our three primary research questions, as they offer wide coverage across boroughs and complaint types. The sufficient volume and temporal details allow us to identify complaint trends and seasonal patterns. The textual data in resolution descriptions enables us to conduct linguistic and sentiment analysis. In addition, the paired satisfaction data allows us to directly test hypotheses about how response phrasing correlates with citizen approval. In conclusion, these datasets allow us to approach the research questions using a multidimensional method. By integrating descriptive statistics, regression modeling, natural language processing, and time-series analysis, we can explore how different variables impact public perception of city services. This alignment between data and analytical goals makes the NYC311 system an ideal case for studying public service dissatisfaction through the lens of applied analytics.

## Analytical Techniques & Findings

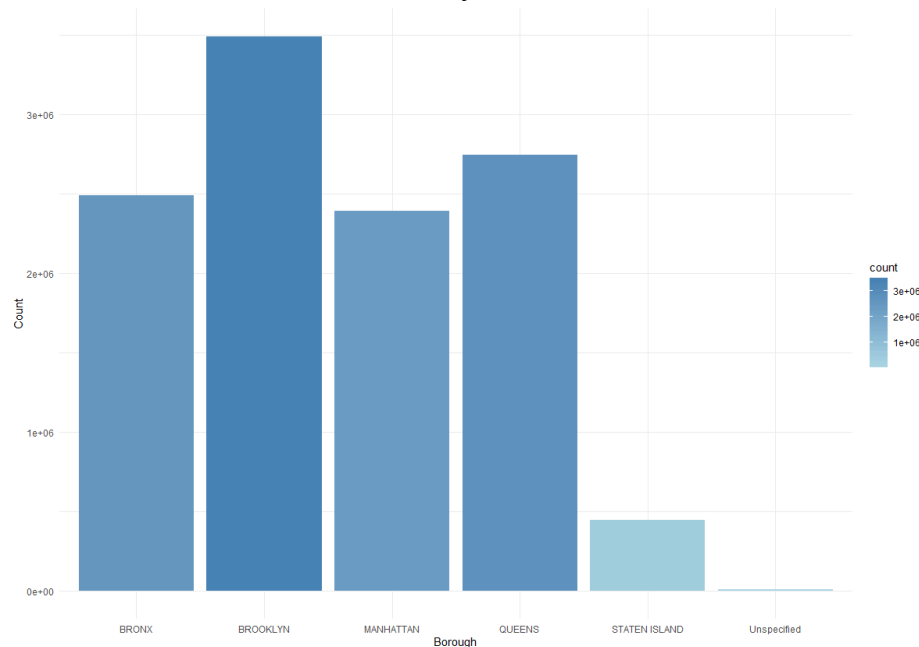
To address the three research questions guiding this project, we applied a multi-method analytical approach that integrates descriptive statistics, natural language processing, and time-series analysis. Each technique was selected to align with the data structure and the specific hypothesis tested within each research question.

*RQ1: What complaint types and borough-level characteristics are associated with low resolution satisfaction or unresolved cases in the 311 system?*

To investigate this research question, our team decided to begin with descriptive statistics and visualizations to summarize the distribution of complaint types, boroughs, and resolution satisfaction levels. We used frequency plots and cross-tabulations to highlight the common complaints and identify boroughs with disproportionately high dissatisfaction. Next, we applied a binary logistic regression model, with satisfaction outcome (Satisfied vs. Dissatisfied) as the dependent variable. Independent variables included complaint type, borough, and agency. This

approach enabled us to quantify the likelihood of dissatisfaction across different categories and geographic areas, which offers insight into structural disparities within the NYC311 service system. In addition, we calculated the odds ratios and confidence intervals to interpret the effect sizes and identify the strongest predictors of dissatisfaction. These findings provided a statistical basis for identifying which issue types or regions may require targeted service interventions.

An examination of NYC's 311 service request data shows apparent differences in the scale and nature of complaints across the boroughs. As seen in **Figure 1**, Brooklyn leads by a wide margin in total complaints, followed by Queens, the Bronx, and Manhattan, while Staten Island has the smallest share. Only a negligible number of cases are recorded without a borough designation. These differences likely reflect not just population size and density, but also differences in infrastructure and in how residents use the 311 system.



*Figure 1. Number of 311 Complaints by Borough*

Patterns in complaint type distribution help explain these borough-level differences. **Figure 2** shows that illegal parking reports are widespread in Brooklyn and Queens, hinting at enforcement challenges in dense, car-heavy neighborhoods. Noise complaints, unsanitary conditions, and water-related maintenance are more evenly distributed across the city, pointing to problems that cut across borough boundaries.

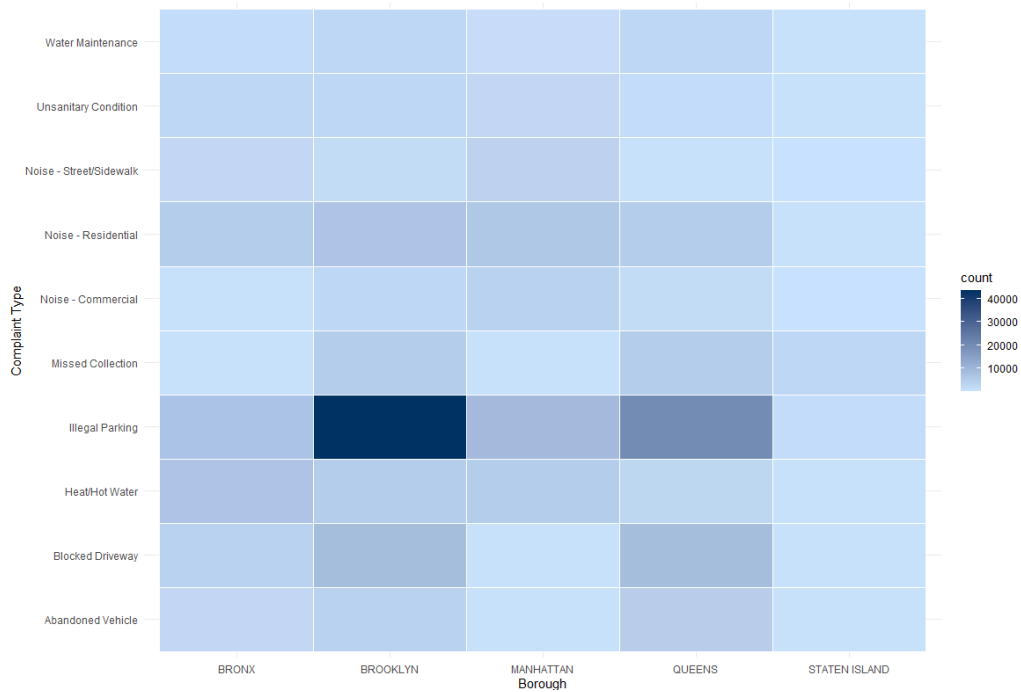


Figure 2. Top Complaints Types Across Boroughs

When we focus on dissatisfaction, the picture changes depending on how it is measured. In **Figure 3**, which uses logistic regression to calculate odds ratios, “Missed Collection” and “Dead Animal” stand out as the complaints most likely to generate dissatisfaction compared to other types. These results show the *relative likelihood* of a negative rating, not the proportion of cases that end in dissatisfaction. Complaints about “Lead” and “Lost Property” also score high, suggesting that sensitive or urgent matters are more prone to adverse outcomes.

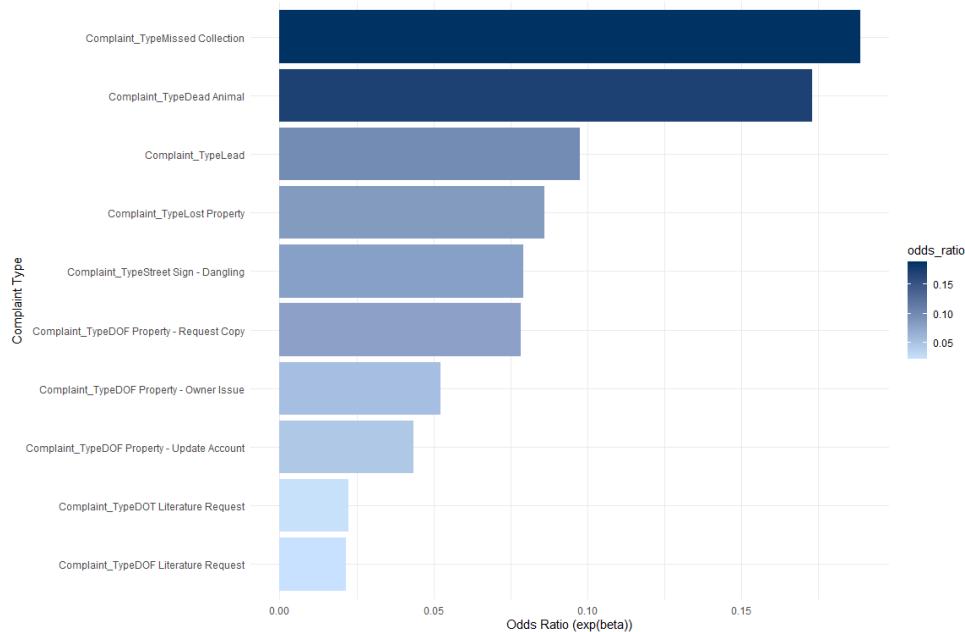
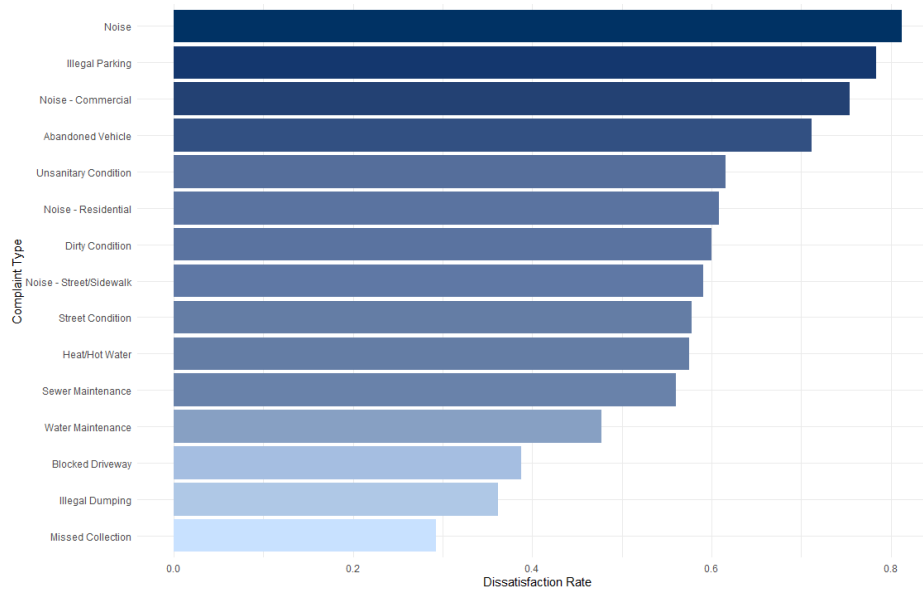


Figure 3. Top 10 Complaint Types with Dissatisfaction

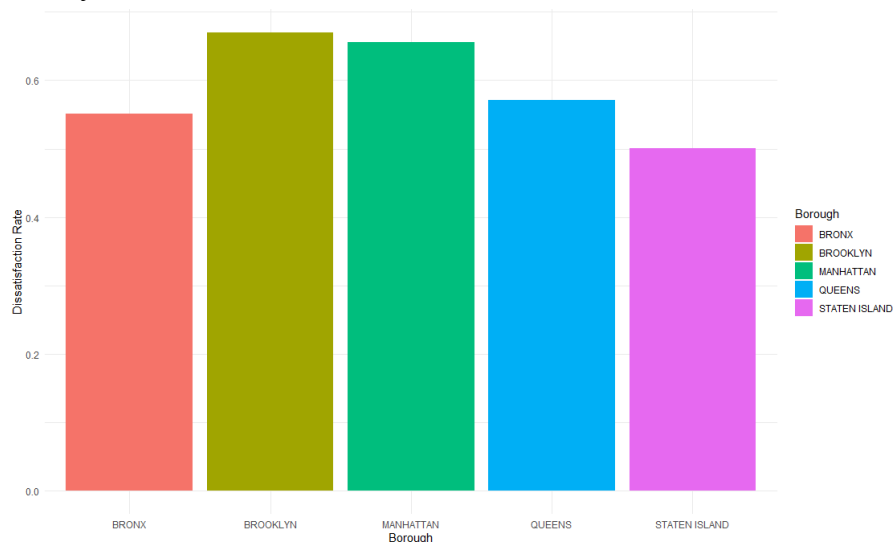


Looking at dissatisfaction as a percentage within each complaint type (**Figure 4**) produces a different ranking. Noise-related issues and illegal parking top the list, both with dissatisfaction rates above 80%. “Abandoned Vehicle” and “Unsanitary Condition” also rank high. These tend to be long-standing urban problems that are difficult to fully resolve, which may explain the consistently poor feedback they receive. In other words, some complaint types have a high probability of causing dissatisfaction for a small group of residents, while others tend to leave *most* complainants unhappy.



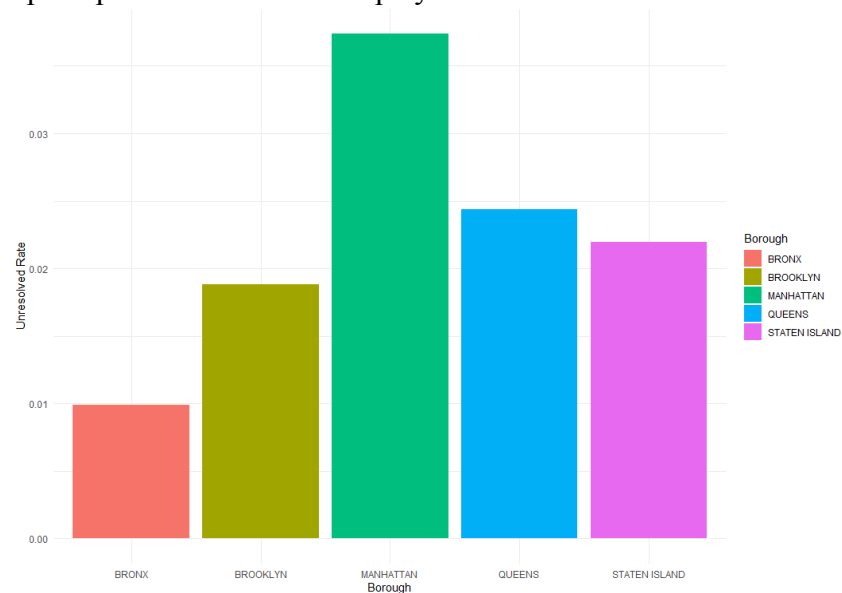
*Figure 4. Top 15 Complaint Types by Dissatisfaction Rate*

Borough-level patterns add another layer to the story. As shown in **Figure 5**, Brooklyn and Manhattan have the highest dissatisfaction rates, both above 65%, while Staten Island and the Bronx are noticeably lower.



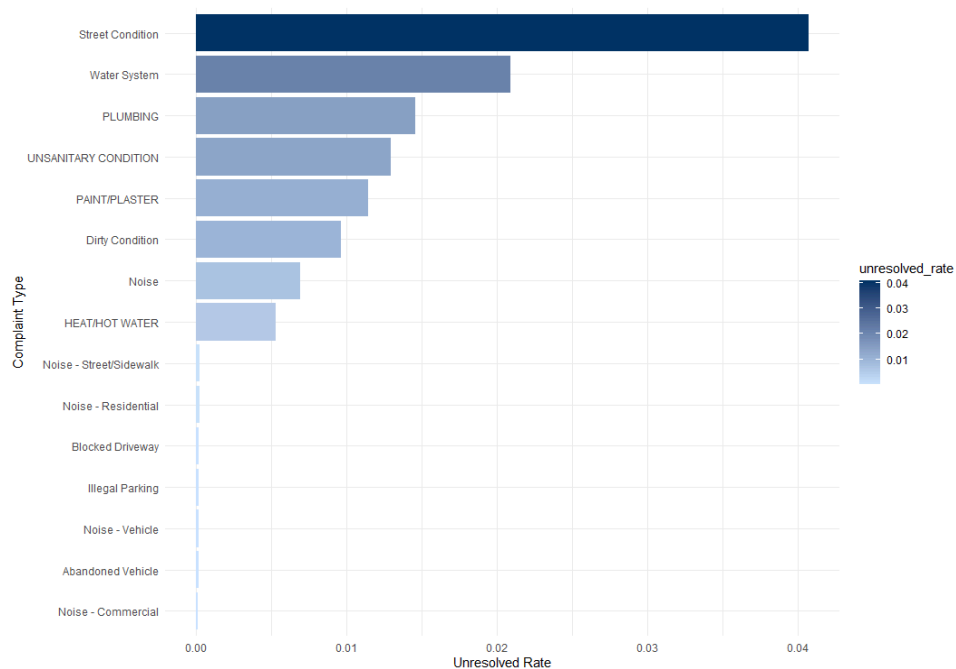
*Figure 5. Dissatisfaction Rate by Borough*

But unresolved cases follow a different pattern (**Figure 6**): Manhattan has the highest proportion of unresolved complaints, followed by Queens and Staten Island, with the Bronx reporting the lowest. This suggests that dissatisfaction is not driven by unresolved cases alone—resident expectations and perceptions of service also play a role.



*Figure 6. Unresolved Rate by Borough*

When we turn to unresolved cases by complaint type (**Figure 7**), infrastructure issues dominate. “Street Condition” leads with over 4% of cases still open, followed by “Water System” and “Plumbing.” These often require complex repairs and coordination between agencies, which can slow down resolution.



*Figure 7. Top 15 Complaint Types by Unresolved Rate*

At the extreme, a small group of complaint types—“Construction Lead Dust,” “Green Taxi Report,” and “Harboring Bees/Wasps”—have a 100% unresolved rate . These outliers are likely the result of very small case counts, unclear jurisdiction, or cases that remain open for extended periods without formal closure.

<b>Complaint_Type</b>	<b>Borough</b>	<b>total</b>	<b>unresolved</b>	<b>unresolved_rate</b>
Construction Lead Dust	BRONX	844	844	1
Construction Lead Dust	BROOKLYN	1217	1217	1
Construction Lead Dust	MANHATTAN	1187	1187	1
Construction Lead Dust	QUEENS	480	480	1
Construction Lead Dust	STATEN ISLAND	36	36	1
For Hire Vehicle Report	Unspecified	75	75	1
Green Taxi Report	BROOKLYN	30	30	1
Green Taxi Report	MANHATTAN	22	22	1
Green Taxi Report	QUEENS	65	65	1
Harboring Bees/Wasps	BRONX	248	248	1

Taken together, the findings suggest that both the type of complaint and borough-specific factors shape dissatisfaction and resolution outcomes in the 311 system. Chronic, quality-of-life issues like noise and illegal parking are the biggest drivers of dissatisfaction, while infrastructure complaints are the most likely to remain unresolved. Brooklyn and Manhattan stand out for high dissatisfaction rates, and Manhattan also struggles with a high share of unresolved cases.

## *RQ2: How are resolution description language patterns associated with citizen satisfaction, and do these patterns differ by complaint type or responding agency?*

In order to look into the relationship between resolution language and citizen satisfaction, we employed a multi-step text mining approach incorporating both qualitative inspection and statistical modeling. The objective was to identify vague or boilerplate language in resolution descriptions and evaluate their association with lower satisfaction levels across agencies and complaint types.

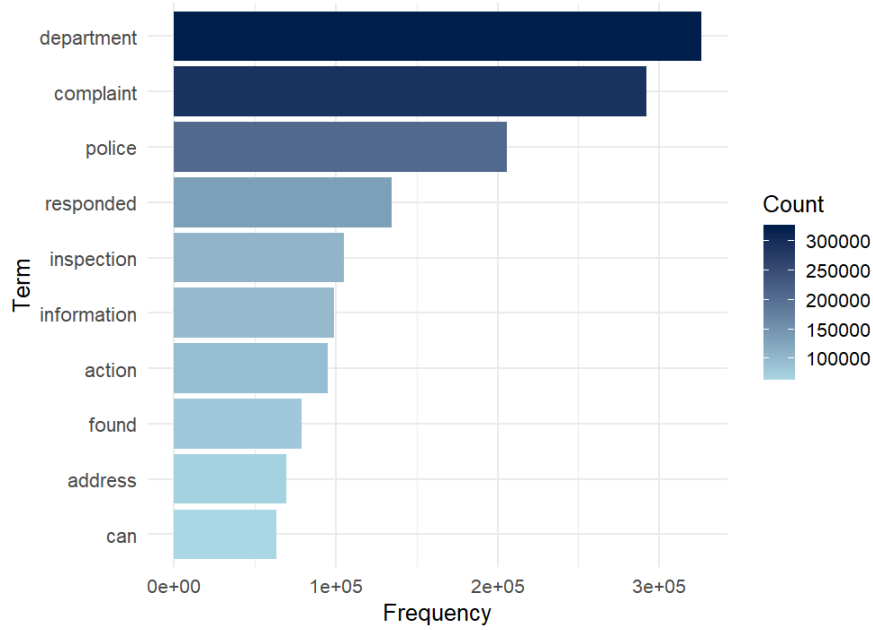
### ***Text Preprocessing and Vague Phrase Detection***

We started by detecting instances of known boilerplate phrases—such as “*no violation observed*”, “*attempted to inspect*”, and “*unable to locate*”—using a regular expression search. These patterns were compiled into a binary indicator (`has_vague_phrase`) to quantify vague responses across records. Moreover, we also applied standard language preprocessing, this process includes lowercasing all text, removing punctuation and numbers, filtering out English stopwords, and applying stemming (using the SnowballC stemmer). This cleaned text served as the input for topic modeling.

Following the preprocessing of the resolution descriptions, we conducted a frequency analysis to identify the most used terms across all 311 service cases. As visualized in the bar chart above, frequently occurring terms such as “*department*,” “*complaint*,” and “*police*” appeared prominently in the resolution narratives. This pattern suggests that resolution language often includes organizational references and procedural responses, which may reflect generalized or standardized messaging by agencies. The high occurrence of such terms raises questions about their informativeness and variability, motivating the need to detect broader thematic structures.

### ***Topic Modelling with LDA***

To uncover recurring semantic patterns in the resolution narratives and find deeper structures in agency responses, we applied Latent Dirichlet Allocation (LDA) model — a generative probabilistic model widely used for topic extraction in textual form. This method assumes that each document, in this case, a resolution description is a mixture of topics, and each topic is a distribution over words. Then, we aggregated token counts by resolution case (`Unique_ID`) and converted the result into a Document-Term Matrix (DTM). In this matrix, each row represents a resolution description, each column a unique term, and each cell the term frequency in that document. This matrix format is essential for input into the LDA model.



*Figure 8. Frequency Analysis Based on Resolution Description*

We applied the LDA() function with  $k = 5$  to extract five latent topics. This number was selected based on a balance between model interpretability and semantic coherence during exploratory runs. Fewer topics (e.g.,  $k = 3$ ) tended to merge distinct complaint categories, while larger values (e.g.,  $k = 10+$ ) introduced noise and redundancy across topics. Using  $k = 5$  allowed us to capture diverse but distinct resolution themes without overfitting or fragmenting the meaning. The model was run with a fixed seed (1234) to ensure reproducibility.

## Top Terms per Topic

Each topic captures a recurring theme in resolution language

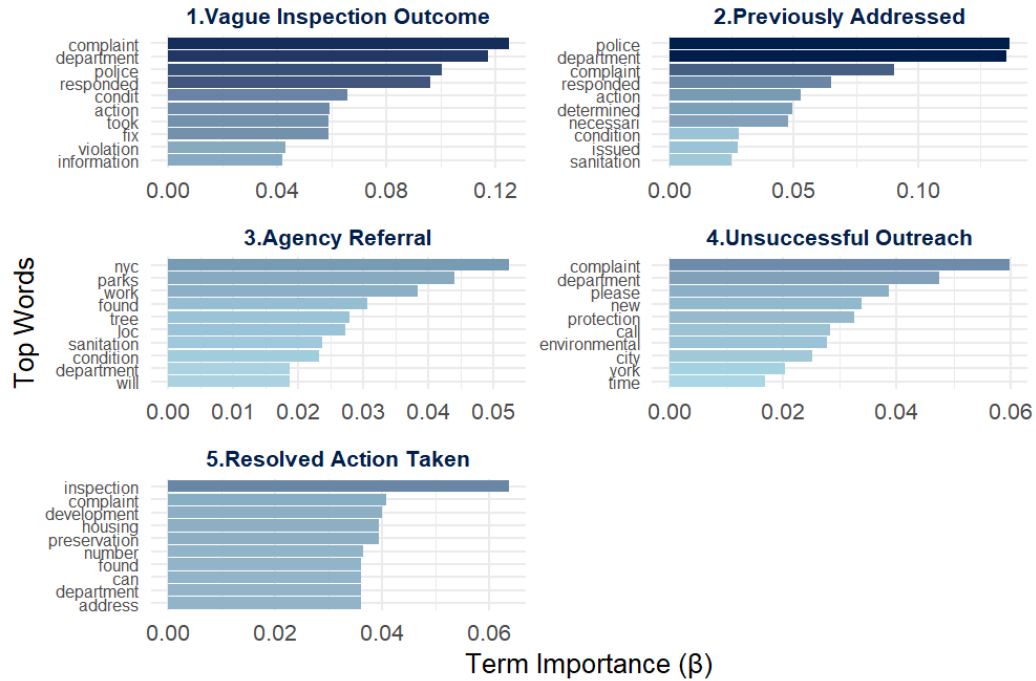


Figure 9. Distribution of Top Terms For Each Topic

Once the LDA model was trained with 5 topics, we extracted the top 10 terms with the highest probability (beta values) within each topic. These terms represent the most semantically dominant words used in resolution narratives associated with each topic. This extraction provides insight into the underlying themes conveyed by city agencies in their written responses.

Each topic was then reviewed and assigned a qualitative label based on the content and context of its top terms. These were derived based on the frequency and importance of terms and linguistic meaning in the context of 311 complaints. This thematic interpretation enabled a more meaningful discussion about how agencies construct their resolution language. Below is a summary of the topics derived:

Topic	Top Terms	Thematic Label	Interpretation
1	unable, access, locate, verify, prior	Vague Inspection Outcome	Describes situations where inspectors could not verify or access the reported issue.
2	duplicate, already, resolved, found	Previously Addressed	Indicates complaints already addressed or duplicated cases.

3	department, forwarded, agency, unit	Agency Referral	Refers to cases redirected to another department or internal unit.
4	attempted, contact, no response, door	Unsuccessful Outreach	Highlights failed efforts to reach the complainant or responsible party.
5	corrected, repair, completed, fixed	Resolved Action Taken	Indicates tangible corrective actions, typically satisfying for residents.

Figure 10. Thematic Label & Interpretation for Topic 1-5

Following the LDA topic modeling and thematic interpretation, we closely examined the semantic content of each topic to identify which resolution styles aligned with vague, non-actionable language.

Topics 1 and 4 emerged as the most representative of this linguistic pattern based on both their top term composition and empirical associations with lower satisfaction scores.

#### **Topic 1: Vague Inspection Outcome**

This topic was dominated by words such as "unable," "access," "locate," "verify," and "prior." These terms typically appear in cases where inspectors could not validate the complaint or complete their task. For instance, a resolution might read: "Attempted to inspect the location, but no violation observed" or "Unable to verify the issue upon arrival."

Such language is characterized by a lack of action or follow-up, ambiguity about the outcome, and passivity in enforcement.

#### **Topic 4: Unsuccessful Outreach**

The top terms included "attempted," "contact," "door," "response," and "no answer." These frequently occur in complaints where agencies claimed to attempt communication or site visits but were unsuccessful. For example: "Attempted contact with resident, no response" or "Left notice on door; no reply received." This pattern signals an incomplete service loop, no confirmation of issue resolution, and deferral of responsibility to unresponsive parties.

What sets these two topics apart is the lack of definitive resolution or remedial action, as compared to Topic 5 (Resolved Action Taken), which included active terms like "corrected," "completed," and "fixed."

#### **Assigning Dominant Topic to Each Resolution Case**

Once the LDA model was trained and topic-term distributions were extracted, we proceeded to assign a dominant topic to each resolution description (document). In LDA, each document is

represented as a mixture of topics, with each topic assigned a probability score (gamma value) indicating how strongly it is expressed in that document.

To simplify interpretation and enable downstream analysis (e.g., regression modeling, satisfaction correlation), we assigned each case to the single most dominant topic — the one with the highest gamma score for that resolution. This approach preserves the most salient semantic category of the case while reducing modeling complexity.

The following steps were as follows:

- We used the gamma matrix output from the LDA model, which contains topic probabilities for each document.
- For each resolution case (Unique\_ID), we identified the topic with the highest gamma value — that is, the topic most representative of the document's content.
- This dominant topic label was then merged back into the original survey dataset using a unique case identifier, so that every complaint record could be associated with a specific topic.

This step allowed us to treat topic assignment as a categorical variable, which could then be used to:

- Analyze topic prevalence across satisfaction responses (e.g., comparing which topics are common in negative feedback).
- Run predictive models to test whether topic types (e.g., vague vs. action-oriented) predict satisfaction.
- Perform per-agency profiling to detect which departments rely more heavily on vague or non-conclusive language.

Assigning dominant topics essentially transformed unstructured text into a structured thematic label, facilitating comparative statistical analysis while retaining interpretability.

### **Labeling Topics 1 and 4 as 'Vague'**

Upon manual inspection of the topic-term distributions, Topic 1 and Topic 4 were characterized by language associated with ambiguity, lack of action, and unsuccessful engagement. Terms such as "unable," "verify," "no access" (Topic 1) and "attempted," "contact," "no response" (Topic 4) suggested resolution processes where the agency failed to complete inspection or connect with the complainant. Based on these patterns, we defined a binary variable `topic_vague` to capture whether a resolution case was assigned to one of these vague topics:

### **Creating a Satisfaction Binary Outcome**

To conduct a predictive analysis, we transformed the survey responses into a binary variable. Specifically, we coded “Agree” and “Strongly Agree” type of responses as 1 (Satisfied) and all others including “Neutral” — as 0 (Not Satisfied). This approach treats satisfaction as a binary variable and assumes a conservative classification: only affirmative responses are treated as true satisfaction. Neutral responses are grouped with negative responses under the “not satisfied” category.



This treatment is acceptable for a binary logistic regression where the objective is to model factors associated with positive citizen sentiment. By treating neutral responses as dissatisfaction, the model could emphasize what contributes to strong satisfaction, not just the absence of dissatisfaction. It ensures a consistent dichotomous outcome variable for logistic modeling and reflects real-world expectations where neutral feedback is often treated as a lack of positive impact. This simplified framework enabled the use of logistic regression, where the dependent variable reflects the likelihood of a satisfactory resolution from the citizen's perspective.

### **Logistic Regression Model: Predicting Satisfaction from Language**

To quantitatively assess whether vague or boilerplate language in resolution descriptions influences public satisfaction, we employed a **logistic regression model**. The binary outcome variable satisfied was coded as 1 for “Yes” and 0 for “No”, reflecting whether the citizen expressed satisfaction with the service resolution.

The predictors included in the model were:

- **has\_vague\_phrase**: a binary indicator for whether the resolution description contained a vague phrase (e.g., "no violation observed", "unable to locate").
- **topic\_vague**: a binary indicator for whether the resolution was categorized under Topic 1 or 4 — topics labeled as vague based on the LDA analysis.
- **Agency\_Name** and **Complaint\_Type**: Categorical controls to account for variation across institutions and complaint categories.

Predictor	Coefficient ( $\beta$ )	Odds Ratio ( $\exp(\beta)$ )	p-value	Interpretation
<b>has_vague_phrase</b>	-0.88	0.42	< 0.001	Cases with vague phrases are 58.3% less likely to receive a satisfied response.
<b>topic_vague</b>	-0.53	0.59	< 0.001	Resolution text under vague topics reduces odds of satisfaction by 41%.

*Figure 11. Table of Logistic Regression Result*

These results strongly support our hypothesis: vague and non-actionable language in resolution narratives significantly reduces the likelihood of public satisfaction.

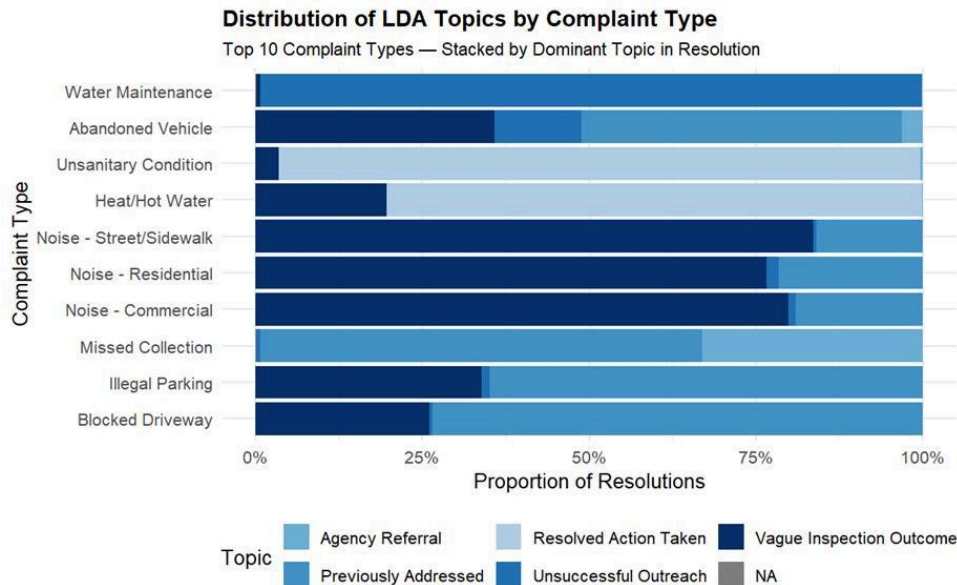
- The p-value for **has\_vague\_phrase** indicates overwhelming evidence that vague language is associated with lower satisfaction.
- The p-value for **topic\_vague** confirms that cases associated with vague LDA topics are also significantly less satisfying.

- The odds ratio for has\_vague\_phrase (0.42) suggests that complaints with vague wording are less than half as likely to result in citizen satisfaction.

Both predictors are statistically significant at the  $p < 0.001$  level, and the odds ratios. This result reinforces the hypothesis: resolution clarity matters. The usage of vague, non-committal language is linked to dissatisfaction.

### **Resolution Patterns by Complaint Type/Agency**

The stacked bar chart reveals notable variation in dominant resolution language patterns across the ten most frequent NYC 311 complaint types. These findings provide insight into how city agencies respond—verbally and operationally—to different issues, shedding light on institutional practices and communicative tendencies.



*Figure 12. LDA Analysis Based on Complaint Type*

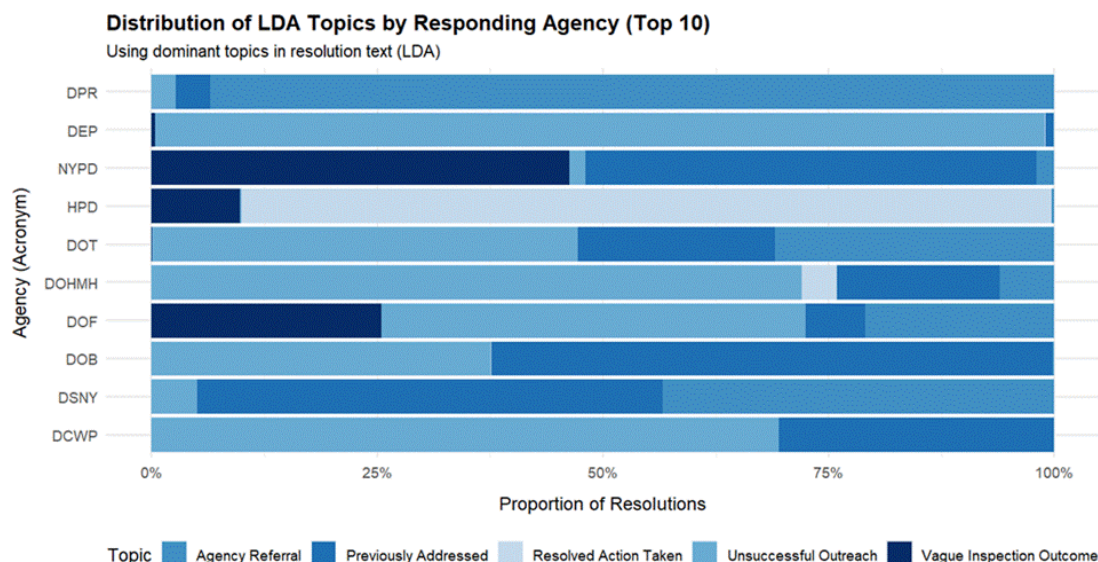
The LDA topic distribution across the top 10 complaint types reveals substantial differences in resolution language patterns based on the nature of the issue:

- Noise complaints (Residential, Commercial, Street/Sidewalk) are heavily dominated by Topic 1: Vague Inspection Outcome, indicating that inspectors often could not verify the issue. This reflects the transient nature of noise and the inherent difficulty in documenting violations after the fact. These categories also show low usage of Topic 5: Resolved Action Taken, suggesting a lack of conclusive follow-up or enforcement.
- In contrast, tangible service complaints such as Missed Collection, Unsanitary Condition, and Heat/Hot Water have a high share of Topic 5: Resolved Action Taken, showing that agencies are able to observe, verify, and resolve these issues more directly. This is

especially evident in Missed Collection cases, which often conclude with clear, documented resolutions like garbage pickup.

- Complaints involving vehicles, including Abandoned Vehicle and Blocked Driveway, show elevated levels of Topic 4: Unsuccessful Outreach. These often require locating individuals, which may fail, leading to template-based closure responses without further action.
- Lastly, Water Maintenance and Illegal Parking show higher proportions of Topic 3: Agency Referral, indicating inter-agency handoffs due to overlapping responsibilities or unclear jurisdiction. A modest presence of Topic 2: Previously Addressed also suggests that some of these cases may have been considered duplicates or previously resolved.

The nature of the complaint significantly shapes how agencies write their resolution narratives. Noise-related complaints are more likely to end with vague or inconclusive language, while tangible service issues tend to result in clearer and more action-oriented descriptions. This variability points to the operational challenges and linguistic conventions tied to each complaint category, which may influence citizen satisfaction and perceived responsiveness.



*Figure 13. LDA Analysis Based on Responding Agency*

Based on the stacked bar chart of LDA topic distribution across top 10 NYC agencies, we observe clear variations in resolution language patterns:

- HPD (Housing) and NYPD (Police) exhibit high proportions of "Vague Inspection Outcome" topics, suggesting a reliance on non-specific or unclear resolution phrases.
- DCWP and DSNY are dominated by "Unsuccessful Outreach" and "Previously Addressed", which may indicate repeated or unresolved service attempts.

- Agencies like DOB (Buildings) and DOT (Transportation) have higher shares of "Resolved Action Taken", reflecting more definitive and actionable outcomes.

The resolution language patterns clearly differ by agency. Certain agencies (e.g., HPD, NYPD) disproportionately use vague or generic phrases, while others (e.g., DOB, DOT) are more likely to report concrete actions. These linguistic tendencies may reflect differing operational procedures, service mandates, or communication norms within each agency.

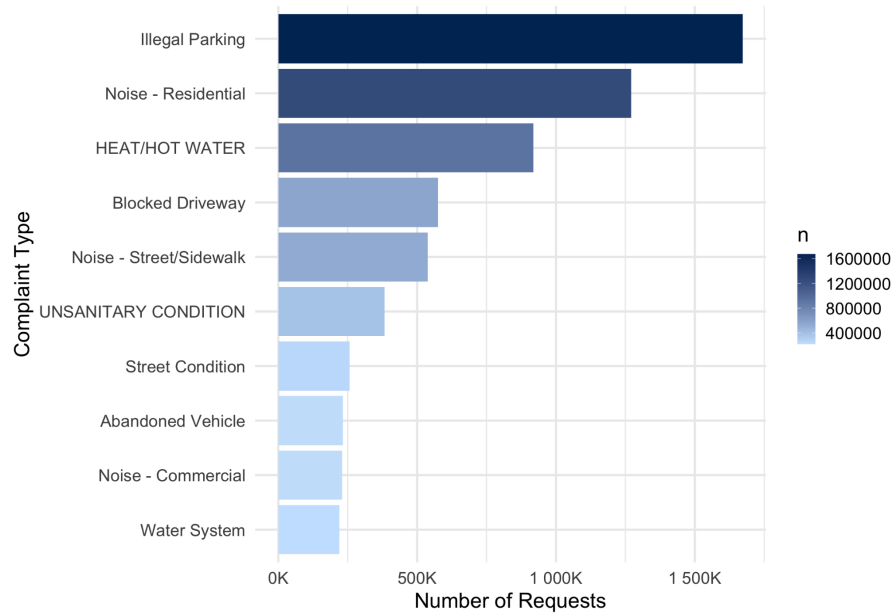
*RQ3: How have complaint volumes for different issues changed over time, and what trends or seasonal patterns can be identified from the 311 request data?*

To explore temporal complaint patterns and assess seasonality, trends, and forecastability of high-frequency NYC 311 service request types, we employed a structured time-series analysis approach. Our methodology involved complaint aggregation, trend-season-residual decomposition (STL), borough-level stratification, and forecasting using two complementary models: ETS and ARIMA. These techniques enabled us to extract both descriptive and predictive insights across complaint categories and boroughs.

### ***Data Preprocessing and Time Series Construction***

We began by filtering the NYC 311 dataset for complete entries with valid "Created\_Date" and "Complaint\_Type" values. Monthly aggregation of complaint volume was performed using floor\_date to standardize complaint timestamps, followed by computation of monthly counts across all complaint types.

Given the extensive variety of complaint categories in the dataset, we focused our analysis on the top 10 most frequent complaint types, which collectively accounted for over 54.38% of all service requests. This decision was based on the observation that, despite the large number of unique complaint types, the top 10 categories alone represented the majority of total volume, making them most impactful for trend analysis and forecasting.



*Figure 14. Top 10 Complaint Types in NYC311*

For each of these top complaint types, we constructed a complete monthly time series object using R's `ts()` function. To ensure temporal consistency, we explicitly included all months within each type's time range, even if no complaints were recorded for some months, by joining with a full sequence of months and filling missing values with zeros. This approach was taken as a precautionary measure to avoid potential issues caused by irregular time intervals, which could interfere with seasonality modeling and forecasting accuracy in ETS and ARIMA.

### ***Visual Exploration and STL Decomposition***

To visually assess the dynamics of complaint volume by category, we first created standard time series line plots using `autoplot` to observe general trends and fluctuations over time. (See Appendix, *Figure 1*) Next, year-over-year seasonal plots (`ggseasonplot`) were used to reveal recurring seasonal patterns in complaint behavior across multiple years for each category. (See Appendix, *Figure 2*)

To gain deeper insight into the underlying structure of the time series, we then applied STL decomposition (See Appendix, *Figure 3*). STL decomposition separates each time series into three interpretable components: Trend, Seasonal, and Remainder. This decomposition allowed us to isolate and analyze meaningful patterns that are often masked in raw plots. By integrating multiple visual techniques, including standard time series plots, seasonal plots, and STL decomposition, we were able to uncover meaningful, complaint-specific temporal patterns that shaped the direction of our subsequent forecasting and policy analysis.

Our exploration of the top 10 most frequent 311 complaint types revealed a wide range of temporal behaviors, including long-term trends, recurring seasonal cycles, and irregular fluctuations. Below, we summarized the most important temporal characteristics observed for

each complaint type.

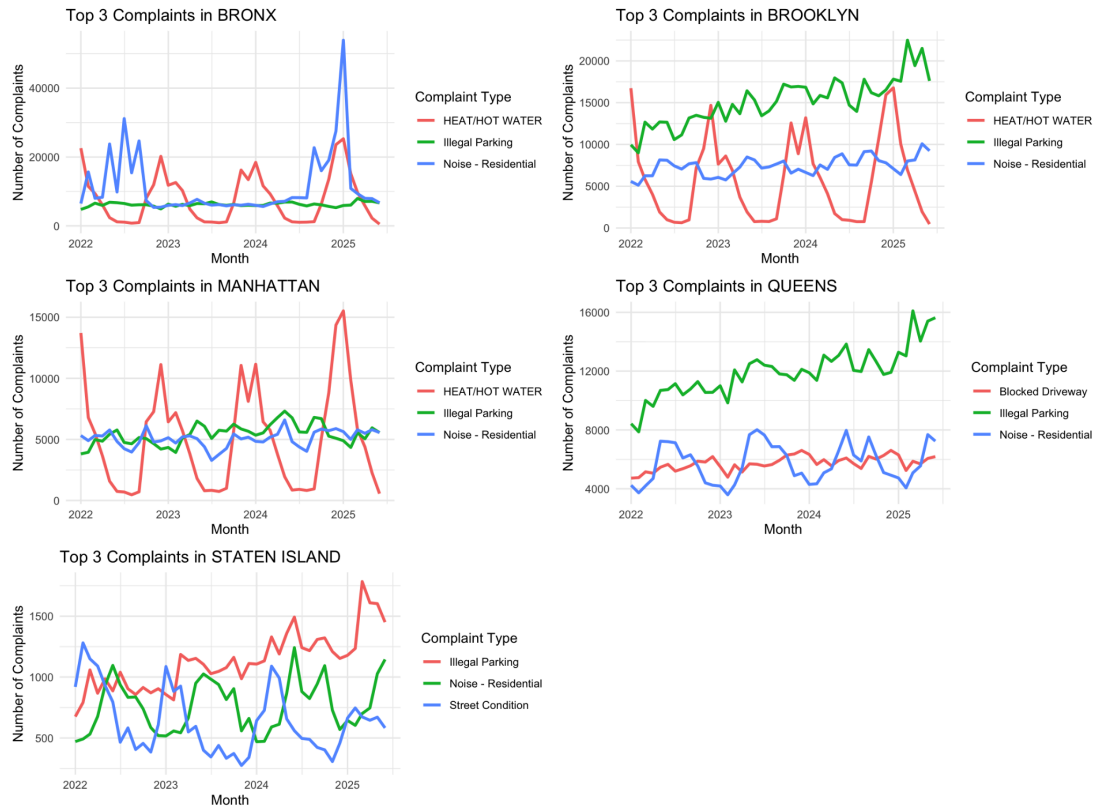
- **Illegal Parking** showed a steady upward trend from 2022 to 2025, indicating persistent or worsening public concerns in those areas.
- **Noise - Residential** exhibited a strong upward trend overall, recovering after a dip in 2023. It also showed modest summer seasonality, reflecting recurring quality-of-life concerns in densely populated neighborhoods.
- **HEAT/HOT WATER** complaints displayed strong winter seasonality, peaking sharply between December and February each year, consistent with seasonal heating needs.
- **Blocked Driveway** showed a gradual upward trend with relatively stable year-round levels and limited seasonality, highlighting a persistently reported access-related issue.
- **Noise - Street/Sidewalk** and **Noise - Commercial** both showed distinct summer peaks, suggesting strong associations with seasonal outdoor activity, nightlife, and tourism patterns.
- **UNSANITARY CONDITION** complaints exhibited moderate summer seasonality and a gradual downward trend, possibly due to improved waste management efforts or reduced reporting during off-peak months.
- **Street Condition** complaints declined gradually over the observed period. A moderate seasonal rise between March and June was also observed, potentially linked to springtime infrastructure wear or roadwork cycles.
- **Abandoned Vehicle** complaints peaked in late 2023 and early 2024, but have since leveled off, hinting at targeted cleanup efforts or shifting public concern.
- **Water System** complaints showed a gradual upward trend and consistent summer seasonality, likely reflecting increased plumbing-related issues and service disruptions during warmer months.

### ***Borough-Level Comparative Analysis***

To investigate how complaint trends differ geographically, we first cleaned the dataset by removing entries with missing, blank, or "Unspecified" values in the Borough field to ensure geographic accuracy.

We then identified the top three most frequent complaint types within each borough based on total counts. Using these top complaint types, we filtered the original dataset to retain only the relevant records and aggregated them into monthly complaint volumes by "Complaint\_Type" and "Borough".

The resulting dataset was used to generate borough-specific time series plots using ggplot(), allowing us to visualize both volume differences and temporal patterns across boroughs. These visualizations revealed location-specific complaint patterns.



*Figure 15. Borough-specific Time Series Plots with Top 3 Complaint type*

In the **Bronx**, complaints related to Noise - Residential showed frequent sharp spikes, especially during winter and summer, likely reflecting chronic disturbances in high-density housing. HEAT/HOT WATER complaints exhibited consistent and strong winter seasonality, peaking reliably each year. Although Illegal Parking was present, it remained comparatively low in volume and stable over time.

**Brooklyn** was dominated by a pronounced upward trend in Illegal Parking, indicating growing congestion or enforcement issues. HEAT/HOT WATER again showed clear seasonal peaks in winter, while Noise - Residential remained relatively flat with minimal seasonal variation.

**Manhattan** followed a similar pattern in heating-related complaints, where HEAT/HOT WATER was by far the most reported issue, with large seasonal peaks each winter. Other issues like Illegal Parking and Noise - Residential remained steady throughout the year, suggesting they are persistent but less urgent concerns in the borough.

In **Queens**, Illegal Parking not only dominated in volume but showed a steep and uninterrupted rise, becoming one of the most pressing issues in the city. Blocked Driveway complaints remained consistently high, reflecting residential car ownership pressures, while Noise - Residential held steady with modest fluctuations.

Finally, **Staten Island** presented a more suburban profile. Illegal Parking continued its upward trend, emerging as the top complaint. Street Condition complaints, while smaller in volume,

showed a visible seasonal peak during the spring months. Meanwhile, Noise - Residential was reported at moderate but consistent levels year-round.

### ***Time Series Forecasting of Key Complaints***

To improve operational planning and enable proactive resource allocation, we conducted time series forecasting on selected high-volume NYC 311 complaint types. Accurate forecasts allow city agencies to anticipate seasonal surges and allocate staff, inspection teams, and budget more efficiently.

To ensure modeling robustness and practical value, five complaint types were selected for time series forecasting based on strong and consistent temporal patterns identified during our prior visual diagnostics. Both **Noise – Commercial** and **Noise – Street/Sidewalk** showed highly regular summer peaks and strong seasonal components, making them ideal candidates for forecasting seasonal surges tied to outdoor activity and nightlife. **Street Condition** exhibited a consistent upward movement during spring and early summer months, reflecting seasonal wear and weather impact, with a clear trend-seasonality structure suitable for maintenance planning. **Blocked Driveway** maintained stable month-to-month behavior with mild seasonality and minimal irregular fluctuations, indicating a reliable base signal for parking enforcement strategies. Finally, **UNSANITARY CONDITION** demonstrated low-noise, steady patterns with minor mid-year seasonality, making it useful for baseline public sanitation resource allocation. Together, these five complaint types combine strong seasonality, trend visibility, and forecasting stability, justifying their selection for more advanced modeling and planning purposes.

For five selected complaint types, we split the time series into training and test sets, reserving the last six months as a holdout period. Two forecasting models were applied:

- **ETS** (Error-Trend-Seasonal): Captures exponential smoothing across trend and seasonal components with automatic model selection via `ets()`.
- **ARIMA** (Auto-Regressive Integrated Moving Average): Configured using `auto.arima()` with exhaustive search to optimize parameters based on AICc.

To evaluate model performance, we visualized the forecasts alongside actual values using `autoplot()`, comparing how well each model captured seasonal cycles and recent trends. Forecast accuracy was quantitatively assessed using standard error metrics.



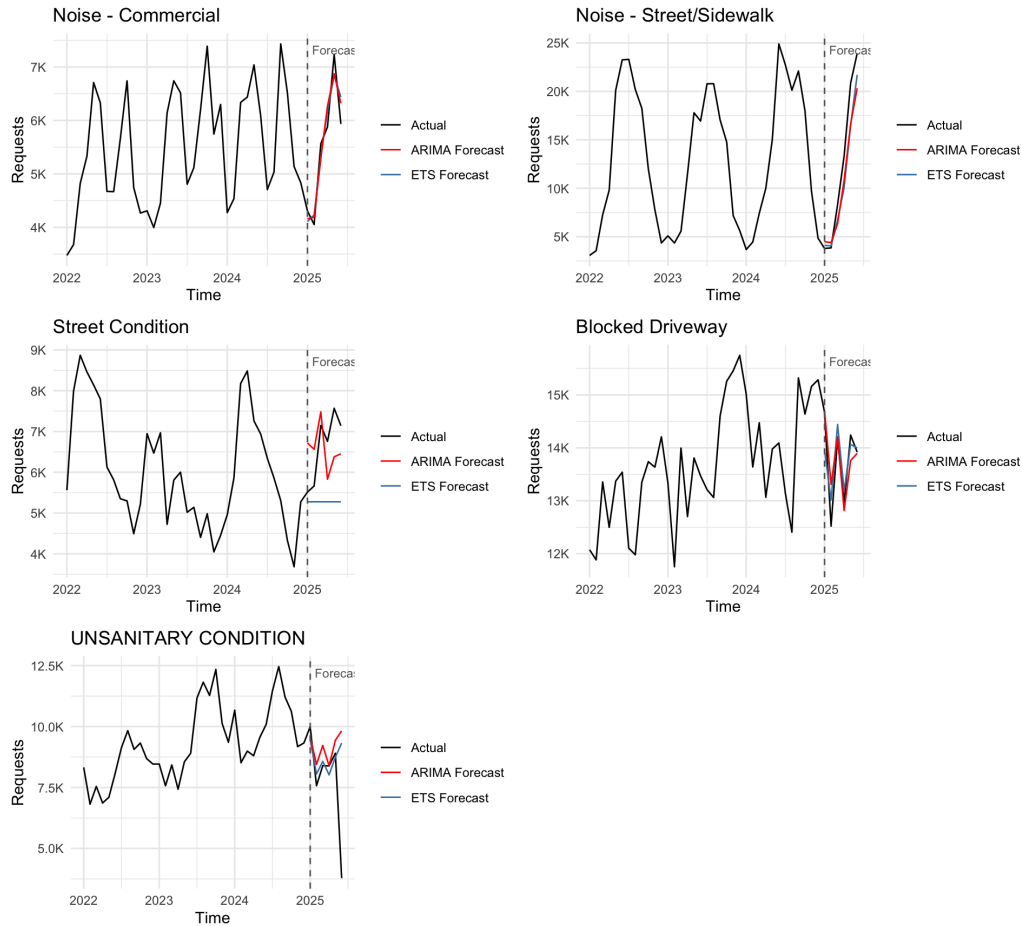


Figure 16. Visualization of Forecast Results

Complaint_Type <chr>	Model <chr>	ME <dbl>	RMSE <dbl>	MAE <dbl>	MAPE <dbl>	MASE <dbl>
1 Noise - Commercial	ETS	-25.9	321.	300.	5.36	0.474
2 Noise - Commercial	ARIMA	-20.9	302.	283.	4.99	0.447
3 Noise - Street/Sidewalk	ETS	1813.	2483.	1989.	14.6	0.892
4 Noise - Street/Sidewalk	ARIMA	1841.	2639.	2257.	18.4	1.01
5 Street Condition	ETS	1357.	1563.	1357.	19.3	1.03
6 Street Condition	ARIMA	61.6	925.	875.	13.6	0.663
7 Blocked Driveway	ETS	-27.4	387.	322.	2.34	0.405
8 Blocked Driveway	ARIMA	-31.3	383.	253.	1.92	0.318
9 UNSANITARY CONDITION	ETS	-857.	2286.	1208.	27.7	1.07
10 UNSANITARY CONDITION	ARIMA	-1309.	2528.	1449.	31.9	1.28

Figure 17. Accuracy Results

The results confirm that ARIMA generally outperformed ETS across most complaint types, particularly in dynamic series. For **Street Condition**, ARIMA demonstrated a substantial RMSE reduction (925 vs. 1563) and lower error magnitudes across all metrics. Similarly, for **Blocked Driveway**, ARIMA achieved noticeably lower MAE (253 vs. 322) and MAPE (1.92% vs. 2.34%). In the case of **Noise – Commercial**, while both models performed well, ARIMA slightly edged out ETS across all error metrics. However, for **Noise – Street/Sidewalk**, ETS achieved

slightly better performance in RMSE and MAE, likely due to its more rigid seasonal structure aligning with the regular yearly pattern. For **UNSANITARY CONDITION**, both models struggled due to higher volatility, but ETS produced marginally better MASE despite higher RMSE. Overall, these metrics reinforce the conclusion that ARIMA offers greater flexibility and predictive accuracy, especially when dealing with complaint types exhibiting evolving trends and moderate noise.

While absolute error metrics like RMSE and MAE can appear large in magnitude, it's important to consider them in light of the **high volume of complaints** these models are predicting. Both ETS and ARIMA forecasts closely align with the actual time series trajectories when viewed in the context of enormous monthly cases. Visual inspection confirms that key seasonal peaks and trends are accurately captured, and forecast deviations represent only a small fraction of overall complaint volume. In practice, this level of performance is **strong and operationally valuable**, providing reliable inputs for resource planning and citizen service optimization.

## Conclusions and Recommendations

Our analysis of NYC311 service requests and surveys, as well as temporal complaint patterns, reveals that dissatisfaction is shaped by both operational and communicative factors. The high-volume and chronic issues, especially noise complaints, illegal parking, and certain sanitation-related concerns, consistently lead to low satisfaction scores, with notable variation across boroughs. Infrastructure-heavy complaints, such as street conditions, face high unresolved rates, which underscore the need for targeted interventions and clearer responsibility assignments.

From a communications perspective, vague or boilerplate resolution language significantly undermines public perception of service quality, even when operational outcomes are similar. Phrases such as “*no violation observed*” or “*attempted contact*” correlates with lower satisfaction scores, which suggests that the clarity, specificity, and perceived accountability in written responses are as critical as the actions taken. On the other hand, from the perspective of temporal trend and forecasting analysis, many complaint categories follow predictable seasonal cycles. These patterns show a clear opportunity for NYC311 to proactively allocate resources and communicate guidance in anticipation of complaint surges. In addition, borough-specific variations highlight the value of tailored strategies that address the distinct service demands of each area.

In conclusion, these insights suggest a multi-pronged strategy for improving citizen satisfaction. To begin with, in terms of operational focus, NYC311 should prioritize persistent, high-friction complaint types, and address bottlenecks in boroughs with elevated dissatisfaction rates. Borough-specific service audits, especially in Brooklyn and Manhattan, can help identify bottlenecks and improve resolution rates. From the perspective of communication standards, NYC311 should retrain staff and revise resolution templates to ensure that language is transparent, actionable, and free of vague phrasing. This is especially important for high-friction complaint types, such as noise or sanitation, which frequently contain non-actionable language. Departments like Housing Preservation and Development (HPD) and New York Police

Department (NYPD) HPD and NYPD, which frequently use vague or inconclusive language, should implement clearer response templates and ensure proper documentation of field verification outcomes. The platform should aim to provide greater clarity in resolutions and strengthen follow-up protocols, including scheduled re-inspections or more responsive outreach.

In addition, seasonal insights should guide staffing and operational planning. For example, additional inspection staff should be allocated for heating-related complaints from December to February in Manhattan and the Bronx, while summer enforcement efforts should focus on reducing noise disturbances. Forecasting tools should be integrated into daily operations, which allows agencies to anticipate surges, balance workloads, and launch proactive public messaging. Tailoring strategies to each borough based on historical and predicted trends, such as prioritizing illegal parking enforcement in Queens or addressing noise and heating concerns in the Bronx, will further enhance responsiveness.

By combining operational focus, clearer communication, seasonal planning, and forecasting-driven strategies, NYC311 can improve public satisfaction, allocate resources more effectively, and build stronger trust between residents and city agencies.

# Datasets Google Drive Link

## Raw Dataset:

311\_Survey\_RAW:

[https://drive.google.com/file/d/1CoNMENOKBH3JEdkazPFoWqtAWb6znIN5/view?usp=drive\\_link](https://drive.google.com/file/d/1CoNMENOKBH3JEdkazPFoWqtAWb6znIN5/view?usp=drive_link)

311\_Requests\_RAW:

[https://drive.google.com/file/d/1q43qd0izeo1xn43T0HsQv4nouFD7bNII/view?usp=drive\\_link](https://drive.google.com/file/d/1q43qd0izeo1xn43T0HsQv4nouFD7bNII/view?usp=drive_link)

## Cleaned Dataset:

311\_Survey: [https://drive.google.com/file/d/1CWOHRqPvBtehXIU\\_xDsS2zct0AHQIVsx/view](https://drive.google.com/file/d/1CWOHRqPvBtehXIU_xDsS2zct0AHQIVsx/view)

311\_Requests: [https://drive.google.com/file/d/1la2nQ3pzu6b3fjfnFSK7tF\\_WeL6cwn7/view](https://drive.google.com/file/d/1la2nQ3pzu6b3fjfnFSK7tF_WeL6cwn7/view)

# References

Citizens Budget Commission. (2025). New Yorkers' ratings of quality of life remain well below pre-pandemic levels.

<https://www.amny.com/news/new-yorkers-quality-of-life-public-safety-ratings-remain-low/#:~:text=Contact%20Us-,New%20Yorkers'%20quality%20of%20life%20and%20public%20safety%20ratings%20remain,improve%20since%202023%2C%20survey%20shows&text=One%20in%20three%20respondents%20reported,good%E2%80%9D%20or%20%E2%80%9Cexcellent.%E2%80%9D.>

DiNapoli, T. P. (2025). DiNapoli releases new tool monitoring NYC 311 complaints. Office of the New York State Comptroller.

<https://www.osc.ny.gov/press/releases/2025/05/dinapoli-releases-new-tool-monitoring-nyc-311-complaints#:~:text=DiNapoli%20launched%20the%20NYC311%20Monitoring,over%20the%20past%20five%20years.>

Hussey, K., & Yan, X. (2025). Data quality challenges in open civic datasets: A case study of NYC311. <https://arxiv.org/abs/2502.08649>.

Justice Data Brief. (2019). Understanding 311 calls for service in New York City. Data Collaborative for Justice at John Jay College.

<https://datacollaborativeforjustice.org/work/low-level-enforcement/justice-data-brief-understanding-new-york-citys-311-data/>.

Kontokosta, C. E., Hong, B., & Johnson, N. (2017). Citizen complaint data as a measure of neighborhood conditions. <https://arxiv.org/abs/1710.02452>.

Lee, J., Chen, M., & Li, W. (2025). Towards an expectation-oriented model of public service quality: A preliminary study of NYC311. ResearchGate. <https://www.researchgate.net/publication/364500301>.

Liu, L., Hong, B., He, C., & McKenzie, G. (2022). Spatial-temporal disparities in resident crowdsourcing delays: Evidence from New York City 311 requests. arXiv preprint arXiv:2204.08620. <https://arxiv.org/abs/2204.08620>.

NYC311. (2025). Resolution satisfaction dashboard. <https://www.nyc.gov/site/311reporting>.

W42ST. (2025, May 31). F is for Frustration: New 311 report cards show NYC agencies failing to satisfy residents. <https://w42st.com/post/f-is-for-frustration-new-311-report-cards-nyc-agencies-failing>.

## Appendix

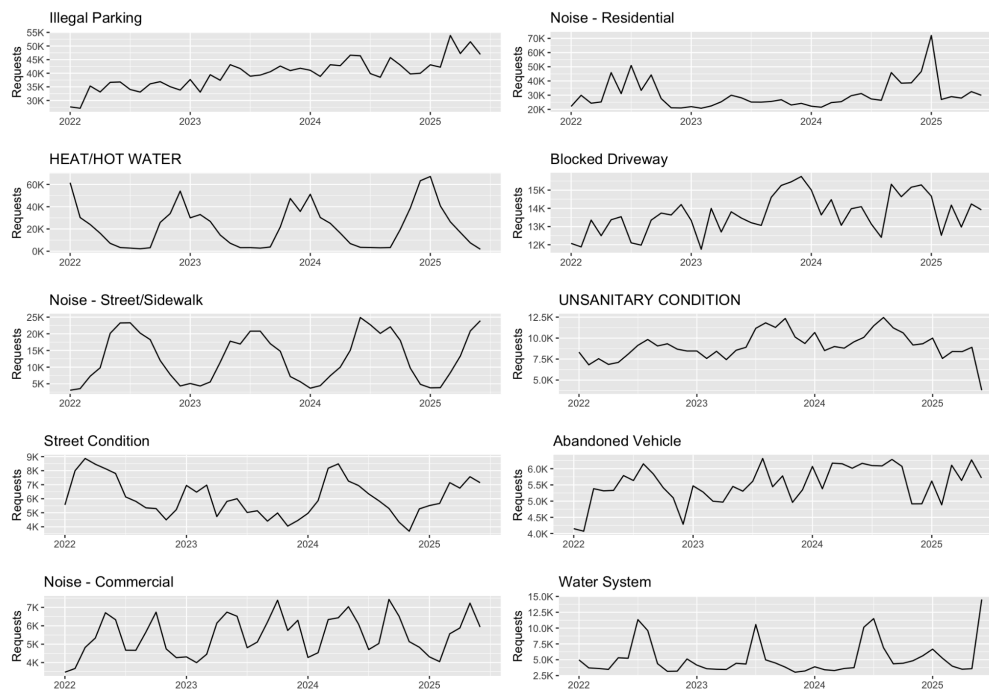


Figure 1. Line plots of monthly volume over time for the top 10 complaint types:

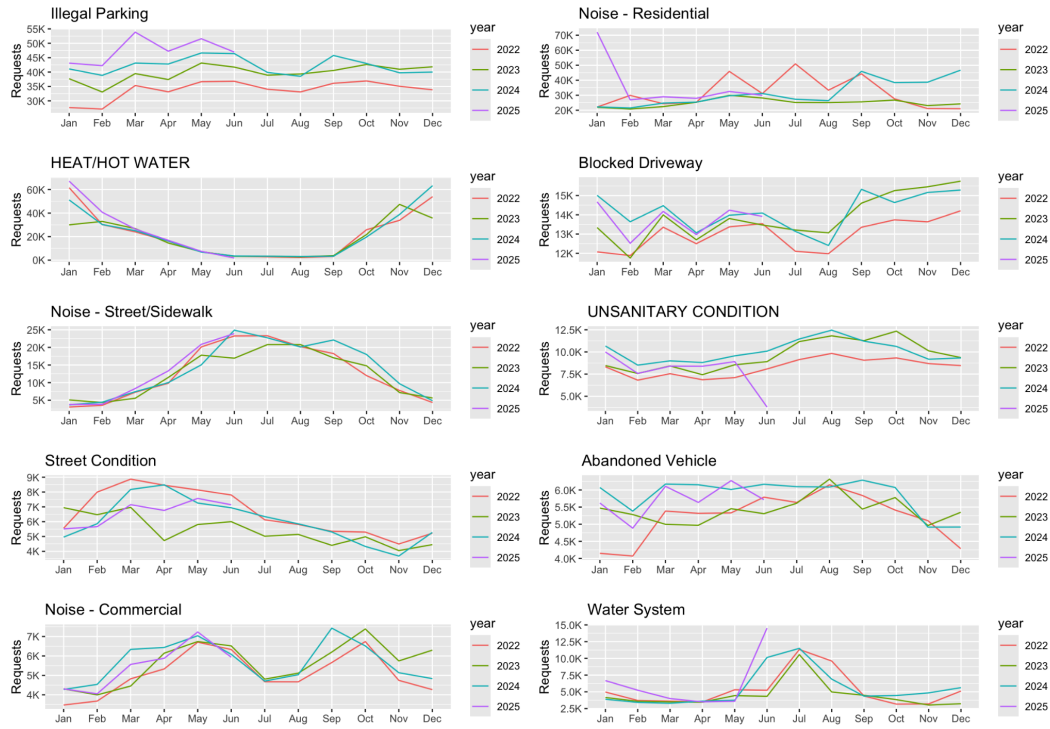
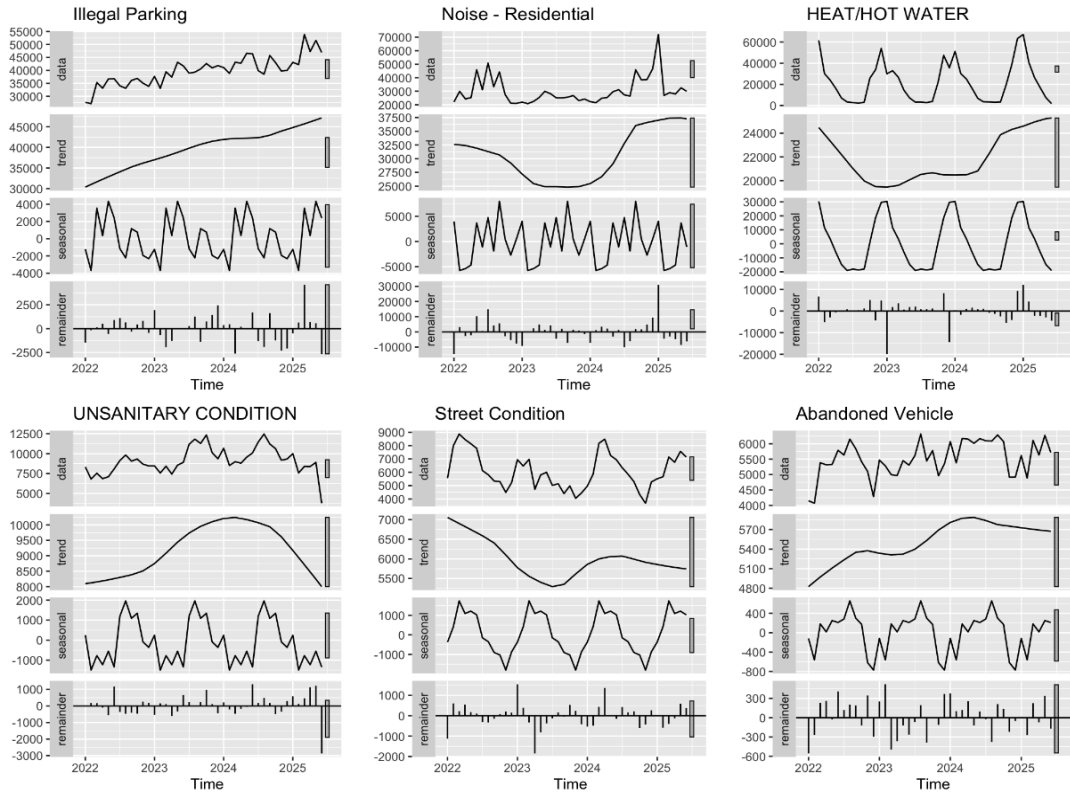
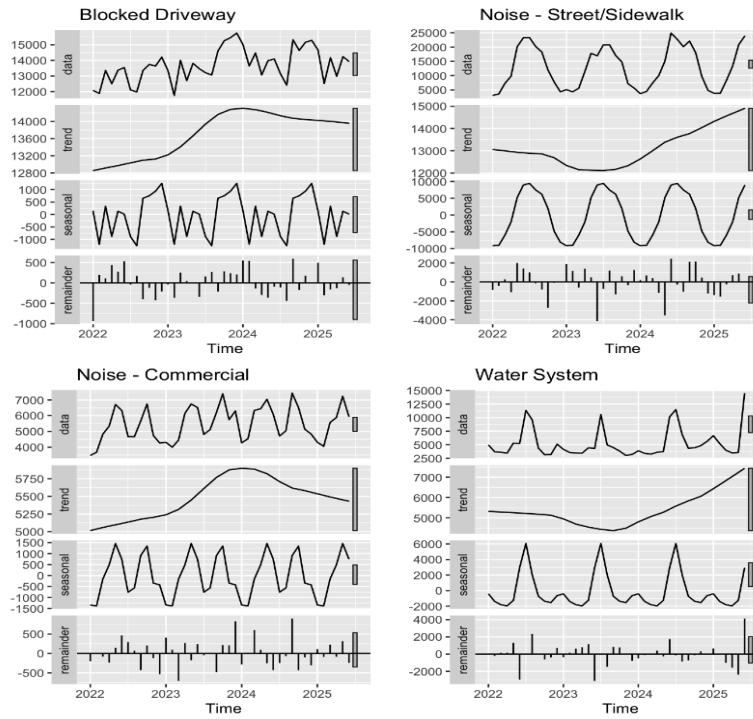


Figure 2. Year-over-year seasonal plots





*Figure 3. STL Decomposition Visualization*