# 311 Complaints in NYC 2010-2021

Habib Alvi, Christopher Lindeman, Vikram Makker, and Saif Qureshi

***Abstract*—An analysis of the complaint data from New York City's 311 hotline. Novel visualization techniques and a demand forecast for noise complaints using time series SARIMA functions.**

***Index Terms*—NYC, 311, complaints, heatmap, time series, regression, forecasting, public administration.**

## I. Introduction

### A. Motivation

BEING able to sift through public data dynamically in a user friendly manner allows residents to be more informed and become engaged with their local government. Also, government agencies can glean insights into their own responsiveness and asset utilization, helping to allocate more of their budgets toward bigger problems. Further, forecasting demand allows the agencies to make informed decisions regarding requests for budget increases when these agencies can show that the proportion of their budget that is allocated for resolving complaints is substandard for the city. Finally, comparing cities against each other can create opportunities for cooperative partnerships in resolving issues that are common, as well as setting performance benchmarks as KPIs to meet.

### B. Problem Definition

Government agencies operate with restrictive budgets, yet must find ways to satisfy its residents. In order to respond to resident demands, agencies must do some forecasting to determine staffing requirements and estimate what budgetary needs will be. Prospective residents will find usefulness in the ability to determine areas where complaints are reported, particularly with regard to choosing their next home. Current residents are able to monitor the status of their neighborhood as well. While the city maintains a Tableau-style dashboard on their site, we were unable to find anything that utilized geospatial representations, and provided this degree of interactivity.

Cities all over the world are have implemented 311 systems, and many of them make their data available to the public, including Alexandria, VA, Köln, Germany, Lamia, Greece, Toronto, ON, Greenwich, UK, and many others. If an analytical framework can solve the problems, or at least provide insights to them, for one of the largest cities in the world, we suspect that it can be ported to other cities.

Authors are students of the OMSA program at Georgia Institute of Technology, Atlanta, GA, 30332.

all team members have contributed similar amount of effort

### C. Survey

Research teams all over North America have critiqued the use of 311 data for policymaking. Other teams have used the data to generate insights in respective, albeit limited, contexts. Attempts have been made to establish correlations between one complaint category and other co-occurrences (Maureen H. Murray, 2018). Books have also been written on increasing citizen engagement towards 311 calls to further improve quality of life, however, they provide little insight on actual analysis (Gavin Newsom, 2014). Analysis of 911 calls (Cramer, 2011) may provide an overarching approach to analysis, however, the size of dataset for this study is exponentially larger. After cleaning the data and performing initial analysis, the team will attempt to investigate if budget allocated to departments had any impact on closure of issues. Similar studies on customer complaints for financial companies show the importance of techniques like regression to improve service levels, however, they fall short of categorizing their complaints effectively (Ayres, 2013). Research shows a spike in calls made to call centers during the pandemic and analysing how those centres process increased volumes can have a direct impact on resident engagement (Andersen, 2021).

Since there is no standard template for logging 311 calls, there is an inherent risk of making faulty deductions when comparing cities (Lingjing Wang, 2017). To mitigate this risk, the team will assess each field and compare like-for-like parameters (Nalchiga, 2017).

## II. Proposed Method

### A. Data Engineering

*1) Dirty Data:* Not only is there a lack of a standard template for logging calls across cities, the amount of mistyped or otherwise poorly logged data requires a bit of massaging to be workable. We decided to throw out many complaint types that were unable to be deciphered, as they were very few when compared with the amount of data points in our data set. Additionally, we collapsed several types of noise complaints into a single 'noise' type to reduce the granularity and focus on the overarching issue of those subcategories all being very similar regarding their response requirements.

*2) Memory Reduction:* The initial complete csv file contained over 26 million data points, in a csv file utilizing over 15GB of storage space. Loading this file into a pandas dataframe used 8.2GB of memory. Many of the features were unimportant to our analyses, which allowed us to reduce the data size significantly. Processing dataframes of this size proved impossible for powerful desktop systems. The following reductions resulted from various processes:

| Process | Memory Reduction |
|---|---|
| using only necessary columns | 80.5% |
| refactor zip codes | 6.3% |
| drop in progress & pending | 13.3% |
| refactor agencies | 15.4% |
| refactor complaints | 11.2% |
| downcast coords | 21.1% |
| Total Processing | 90.4% |

**TABLE I:** Memory Reduction per Feature Engineering Process

Datetime information was reduced to int8 columns for each day, month, and year. These needed to be recombined later, but saved memory during interim processing steps. Ultimately, the result of preprocessing allowed for all members to perform their work locally, with the final size of the csv created from the dataframe after all processing taking up just 1.48 GB of storage and utilizing 514.4 MB of memory - transforming this from a highly demanding cloud-based project into one that can be performed independently by several researchers, as we have done.

For successful interpretation of the new data, we created lookup dictionaries to refer to for the coded complaint types and agencies. These proved essential to interpreting the data throughout our analysis.

## B. Static Visual Analysis

Rendering graphics using matplotlib and seaborn allowed us to visually inspect the data to find general truths and trends, such as learning that noise complaints are significantly the most prevalent, and that Staten Island seems to resolve complaints much faster than the other boroughs.

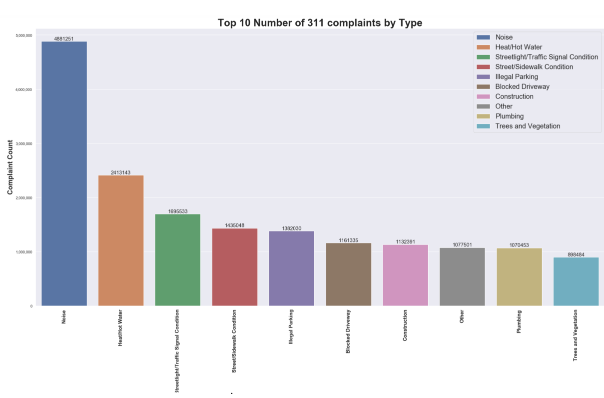


**Fig. 1:** Top 10 Complaints by Type

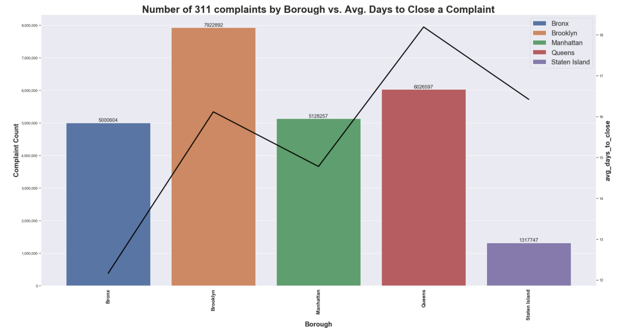


**Fig. 2:** 311 Complaints by Borough vs. Average Days to Close

These graphs work as a quick and dirty way of gaining understanding of simple truths and trends without devoting the significant amount of time that would be needed for a more thorough analysis.

## C. Interactive Heatmap Visual Analysis

Utilizing a random sample, equal to 0.5% of the original data, we were able to deploy a web app using the Dash framework that allows users to visualize the complaint data as a heat map overlaid on a street map of NYC. Users can filter the data by complaint type and/or agency for a range of years selected by a range slider. The range slider can select a single year or the aggregate data of the years in question. An animation is returned which plays through the months of the year, allowing users to see the seasonal changes. After incurring significant charges on GCP, we decided that we could use a small random sampling for the visualization and users would still get a good understanding of the location of complaints, particularly when using a heat map, and chose PythonAnywhere as our host. The Dash app can be viewed here.

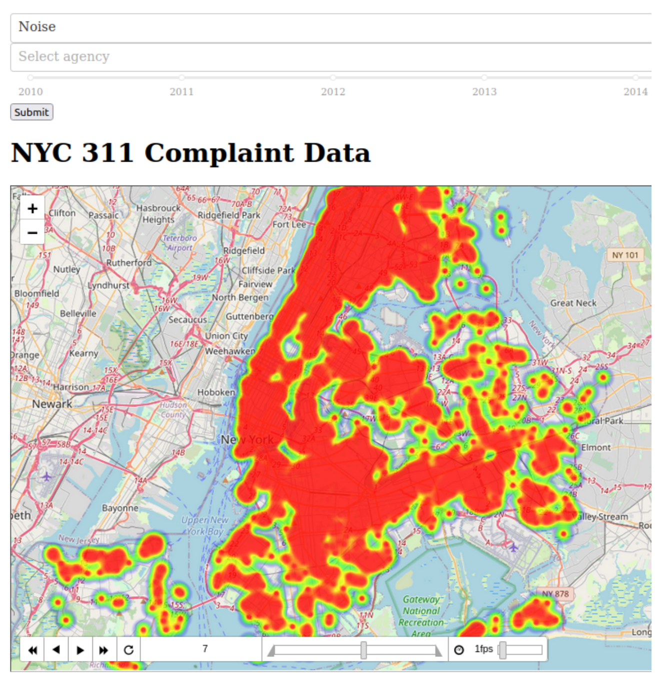


**Fig. 3:** Heatmap of Noise Complaints, July 2020

### D. Budgetary Regression Analysis

The team set out to provide the government agencies with tools to better allocate their funding and resources. Therefore, it was imperative a relationship between budgets, and service levels be established. Service level is defined by complaints closed divided by total complaints logged in a year. The data suggests that New York City has 32 agencies, and since 2011, it logged approximately 27 million complaints. Given the absence of financial information in 311 data, external resources (Department of Budget, 2021), (Department of Budget, 2019), (Division of the Budget, 2018), (New York Building Congress, 2015), (NYC Open Data, 2021), (NYC Open Data, 2021) were found to investigate if operating budgets had any correlation with agency service levels. Since there is little information on how funds are allocated by each agency, it is assumed the entirety of the operating budget was spent on handling 311 complaints. To get a more realistic assessment, it is assumed that the fraction of budget spent on handling a category of 311 complaints is equivalent to the fraction of the said category of complaints processed by an agency. For instance, if 50% of complaints handled by NYPD were noise complaints, then, it is assumed that 50% of the budget was spent on handling noise complaints. We are aware that this approach has its limitations by way of its assumptions.

### E. Complaint Forecast

To predict demand for noise complaints, we chose to implement a seasonal ARIMA model. Our decision to focus solely on noise complaints was due in part to the fact that noise complaints are by far the most common complaint type, and that this can be extrapolated into any other, or even all complaint data. Abandoning a neural network approach due to computational limitations from dummy variable creation (over 10 petabytes) and inability to determine the most relevant factors, we have implemented the statsmodels package. This required refactoring the data into a format wherein data was simply filtered and regrouped by monthly date using a count method. Determining the parameters of the time series model was derived with a loop function that reported the AIC for various models, which allowed us to select the optimal model. After running the model on the complete data set, we noticed that the events of 2020 created anomalies in the forecasting, which is to be expected, so we revisited the approach using data until March, 2020 to compare. The limited data produced a tighter confidence interval, while the complete training data appears to be a more likely estimate as Covid-related restrictions lighten and people return to pre-Covid behavior.

## III. Evaluation

### A. Experiments

*1) Time Series:* We fit the SARIMA model to the time series data on noise complaints on the complete data set as well as a subset of the dataset ending before the covid-19 lock downs began in March 2020, which we expected caused major variation in the data. Comparing the forecast charts, we can see that the confidence interval for the years 2022-2023 is significantly wider than that provided by fitting the model to the data ending with March 2020. However, the model without 2020-2021 data will fail to take into account any fundamental changes in society as a result of the pandemic.

*2) Complaints Analysis by the volume:* We created visualization charts on all the complaints data to understand following aspects

- Borough with highest number of complaints
- Government agency tackling most number of the complaints
- The type of complaint that is most reported.
- The areas where the complaints originated.

*3) Complaints Analysis by the Response Time:* Response time is the best indicator of the performance of the city departments and we analyzed following aspects

- Borough with lowest and highest response time.
- Agencies that take longest to answer the complaints for every borough
- The changes in the response time over the years
- Neighborhood response time heatmaps by every agency
- Sub categorization of the complaints and how it changed over the years

*4) Analysis for Residential Noise Complaints:* As residential noise complaints are the largest complaint type, we dig deeper on that and find out where it is mostly coming from and if the response is adequate for these complaints.

*5) Benchmark Performance of NYC data Against Los Angeles:* One of the most important features was to compare NYC 311 response time against any big city. LA 311 call data was publicly available so we used that to benchmark it against NYC 311 data. Due to the storage and time restraints, we only downloaded a sample of 1.0 million records of year 2020.

*6) Complaint and Budgetary Regression:* A linear regression on closing complaint cases, using noise as a proxy for overall performance, is to be performed using the sci-kit learn library. We used complaints logged by year[a], complaints closed by year[b], and the performance level, indicated by $a/b$. It is to be noted that linear regression was performed on a sample of all available data (pre, and post pandemic).

As seen in Table 2, the linear regression model with the best fit has three predictors, total complaints logged in a year, total complaints closed in a year, and the performance level.

| Created | Closed | Performance | $R^2$ |
|---|---|---|---|
| X | | | 0.57 |
| | X | | 0.57 |
| | | X | 0.21 |
| X | X | | 0.62 |
| X | | X | 0.59 |
| | X | X | 0.59 |
| X | X | X | 0.72 |

**TABLE II:** Linear regression summary for NYPD handling of noise complaints. Response variable: Allocated budget

## B. Outcome

*1) Time Series:* Ultimately, we will have to wait and see which model proves more accurate, though it is the opinion of this team that the model including the complete data set will prove to be more representative of the actual future complaints, as its reversion to the mean will begin from the higher volume that was experienced as a result of 2020 lock downs. When using incomplete training data, the maximum value forecasted for 2022 is roughly 2100 complaints, versus the approximately 3500 forecasted when using the complete data.

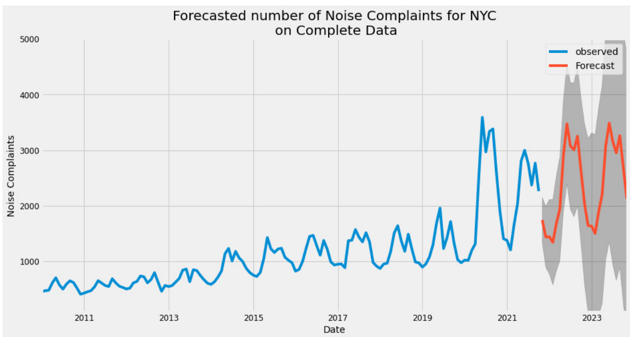


**Fig. 4:** Time series prediction on complete data

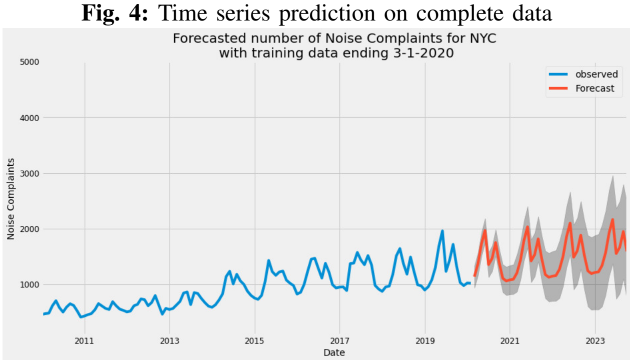


**Fig. 5:** Time series prediction on data ending 1 March, 2020

*2) Complaints Analysis by the call volume:* The outcome from this experiment gave us some very useful insights. Here is the outcome of these experiments.

- Brooklyn has the largest number of complaints, which is not at all surprising considering it is the largest borough with the highest number of residents.
- NYPD and the Department of Housing receives most of the calls. The interesting fact is that 3rd in the list is the Department of Transportation and it does not get half of the calls as the first two.
- Noise is the most reported complaint and the second in the list is Heat/Hot water and it is even less than half of the noise complaints so the city needs to focus more on reducing noise complaints.
- Most of the complaints are reported from residential buildings and it's also not surprising considering there are a lot of smaller buildings and some very bad living conditions in many of those buildings.

*3) Complaints Analysis by the Response Time:* The outcome of this experiment gave us insight on the response time and which agency and neighborhood needs more attention. Here are a few of our observations.

- Brooklyn has the highest complaints but the response time for Queens is longest. Also, the interesting fact is even Staten Island has the least number of complaints, they have the second highest response time.
- The Fire department in Brooklyn, Building Department in Queens, Dept. of Parks and Rec in Bronx and Staten Island and Taxi and Limo Commission in Manhattan are the slowest agencies to respond. Only interesting enough is the Fire department because it needs to answer much quicker.
- Overall response time has been better in the last few years.
- Heatmap gave us insight into agencies that are taking longer. E.g in Staten Island Economic Development Corporation has the slowest response time and is the same as the Department of Education in Brooklyn.

*4) Analysis for Residential Noise Complaints:* This analysis shows some interesting facts. The affluent neighborhoods have the lowest noise complaints but for poor neighborhoods across all boroughs the number of complaints are pretty consistent. Also, the geo map visualization discovers that some of the poor neighborhoods that have the most complaints get the slowest response. This may be attributed to social inequity that exists in many cities, and deserves further investigation.

*5) Benchmark Performance of NYC data again LA:* The LA 311 call data was downloaded from data.lacity.org website and the visualization shows that the agencies in LA usually respond within 10 days whereas NYC takes around 20 to 25 days. Further investigation into the operations and budgets of the agencies would be needed to determine the efficiency of either city with regard to its complaint resolution.

*6) Complaint and Budgetary Regression:* The relatively high accuracy of the model indicates that even during the times of pandemic when complaints were overwhelmingly high (as seen in Fig. 4), NYPD maintained a consistent performance level. Since the model with highest R2 has three predictors, plotting it on a two-dimensional chart would be impossible. However, data has been tabulated for reference in Appendix A. There is a strong correlation between an agency's performance and budget allotment. Therefore, agencies can use the forecast in the years that follow when they request for funding from the local governments, and use it to prioritize and then alleviate the concerns of the residents.

# IV. Conclusion

THE analytical investigation has resulted in several insights that agencies can use to make informed decisions. The analyses have uncovered several simple truths that policymakers can use to better serve their residents, such as preparing for greater noise complaints as summer approaches, and the general trend towards more complaints, likely a result of more people adopting the use of the 311 system as time progresses. Community organizers and activists can begin to identify inequities in the resolution of complaints between areas of affluence and those of economic struggle. Generally, residents can better understand the prevalence of issues in their neighborhoods, and people considering relocation will have an investigative tool at their disposal from the interactive heatmap tool. This tool could be improved significantly by using the complete data set, however there was no way to host that much data with our budget. We have begun to assess the efficacy of NYC by comparing it to LA, which cities can use to compare their performance and identify both shortcomings and successes in their complaint handling. In summation, the analyses performed have opened a door to deeper investigation of a variety of hypotheses that can benefit residents, both current and future, as well as agencies, activists, communities, and budget officials.

Dec 5, 2021

## Appendix A
## Raw data used for regression

| | countAllOpen | countAllClosed | closure | allocated |
|---|---|---|---|---|
| **0** | 8052 | 8051 | 0.999876 | 2.544532e+09 |
| **1** | 9184 | 9170 | 0.998476 | 2.584718e+09 |
| **2** | 10851 | 10846 | 0.999539 | 2.599699e+09 |
| **3** | 14115 | 14114 | 0.999929 | 2.605834e+09 |
| **4** | 16190 | 16188 | 0.999876 | 2.732512e+09 |
| **5** | 17879 | 17873 | 0.999664 | 2.832266e+09 |
| **6** | 18913 | 18912 | 0.999947 | 5.658054e+09 |
| **7** | 18605 | 18605 | 1.000000 | 8.623654e+09 |
| **8** | 20365 | 20365 | 1.000000 | 9.328130e+09 |
| **9** | 37111 | 37111 | 1.000000 | 9.367729e+09 |

## Acknowledgment

The authors would like to thank Professor Polo Chau at Georgia Tech for providing the impetus to form this group and collaborate with each other to learn new technologies and determine new approaches to existing problems.

## References

[1] Andersen, D. (2021, September 10). How to Analyze Call Center Data to Improve Efficiency. Retrieved from Invoca: https://www.invoca.com/blog/how-to-analyze-call-center-data-to-improve-efficiency
[2] Ayres, I. L. (2013). Skeletons in the Database: An Early Analysis of the CFPB's Consumer Complaints. SSRN Electronic Journal.
[3] Cramer, D. &. (2011). Predicting 911 Calls Using Spatial Analysis. Studies in Computational Intelligence, 15-26.
[4] Department of Budget, 2019. Department of Transportation. [Online] Available at: https://www.budget.ny.gov/pubs/archive/fy19/exec/agencies/appropData/TransportationDepartmentof.html
[5] Department of Budget, 2021. Transport, Department Of. [Online] Available at: https://www.budget.ny.gov/pubs/archive/fy21/exec/agencies/appropdata/TransportationDepartmentof.html
[6] Division of the Budget, 2018. Transportation, Department of. [Online] Available at: https://www.budget.ny.gov/pubs/archive/fy18archive/exec/agencyPresentations/appropData/TransportationDepartmentof.html
[7] New York Building Congress, 2015. Capital budget snapshot:New York City Department of Transportation. [Online] Available at: https://www.buildingcongress.com/advocacy-and-reports/reports-and-analysis/infrastructure-update/Capital-budget-snapshot-new-york-city-department-of-transportation.html
[8] NYC Open Data, 2021. Expense Budget. [Online] Available at: https://data.cityofnewyork.us/City-Government/Expense-Budget/mwzb-yiwb
[9] NYC Open Data, 2021. Expense Budget Funcding - All Source. [Online] Available at: https://data.cityofnewyork.us/City-Government/Expense-Budget-Funding-All-Source/39g5-gbp3
[10] Gavin Newsom, L. D. (2014). Citizenville: How to Take the Town Square Digital and Reinvent Government. Penguin Books.
[11] Lingjing Wang, C. Q. (2017, October 17). Structure of 311 service requests as a signature of urban location. Plos One. Retrieved from Plos One: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0186314
[12] Maureen H. Murray, R. F. (2018). Public Complaints Reflect Rat Relative Abundance Across Diverse Urban Neighborhoods. Frontiers in Ecology and Evolution.
[13] Nalchiga, S. &. (2017). Achieving Interoperability of Smart City Data: An Analysis of 311 Data. Journal of Smart Cities.