Analyzing 311 Responses

Final Report

Cheng-Ping Lin (Class of 2023, momolin@bu.edu)
Kyran Chen (Class of 2023, kyc0510@bu.edu)
Weichen Jiang (Class of 2023, weichen@bu.edu)

Introduction

Project Goal and Overview

311 is a telephone number that connects people with service representatives who are always ready to provide non-emergency city services and information. Our project aims to understand the city's response to 311 service requests by analyzing 311 data. By doing this, we are able to understand which communities feel empowered to demand services and how the city responds. Specifically, we collected and pre-processed the data from the CLIMATE READY BOSTON SOCIAL VULNERABILITY dataset and the 311 requests from the past 12 years. We then normalized the data to ensure all variables have fair impacts to the result.

Base Analysis

How did we calculate the social vulnerability index of each census group? That is, how do we define social vulnerability status of a census group?

We used the ranking method to calculate CDC/ATSDR SVI. We first divided the number of people in each social group by the total number of the population. Then we sorted census groups by the rate contributed by each social group, and for every census group at the first 10% in each sort, add one point to their SVI. After going through all social groups, we got a rank of social vulnerability status.

Data cleaning:

When calculating the ratio, we've researched to find that there are some groups with a small amount of request cases, the ratio would be extremely high. Therefore we blocked those ratio to get a more generalized pattern.

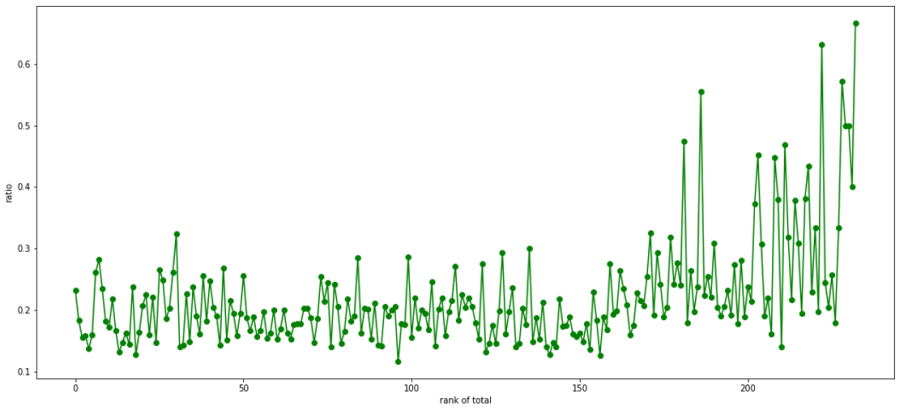


Figure 0 The pattern between population and request ratio

While analyzing the closure rate, we've found a type of request called 'general request' which contains all uncategorized requests. As we need to deal with type in this case, we simply ignore this type.

How do we estimate 311 requests based on the calculated SVI?

With the SVI of all census groups, we tried two methods to analyze how well the SVI predicted the load of 311 requests. For every district in each SVI rank, if the position in 311 cases number list is in the range of it's rank's district, we classify the SVI rank as a good predict (For example, let's say district EX has a SVI rank 2 and all districts in 151 to 171 position has a SVI rank 2. If we found that in 311 cases listed the position of

district EX is between 151 and 171, we say SVI rank was a good predictor). As a result, 135 out of 183 census blocks received a good prediction. We think the bias between the two features might be that there are spam calls, and also the SVI rank only gives an approximate prediction as it is a rank. Below is the visualization of SVI we calculated.

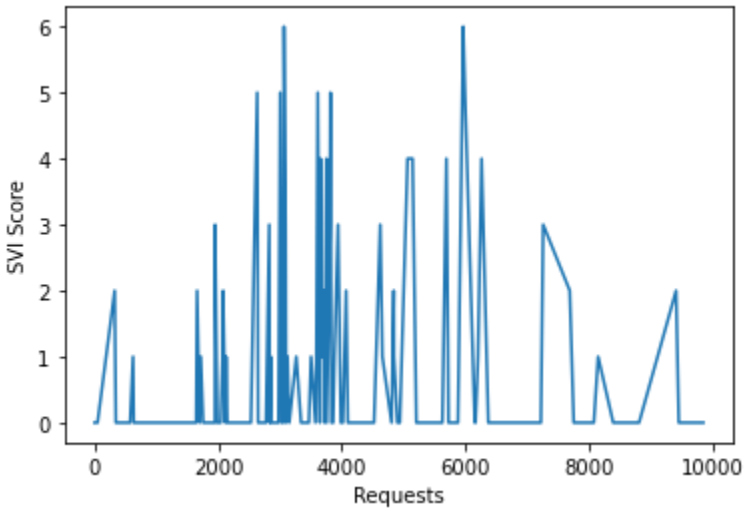


Figure 1. SVI scores at different number of requests

In order to give a better view of the pattern between svi variables and requests, we also do the regression that uses all the variables to fit the request numbers of all 181 mentioned census tracts to calculate how much each variable contributes to the request number. We did the regression with and without extreme cases:

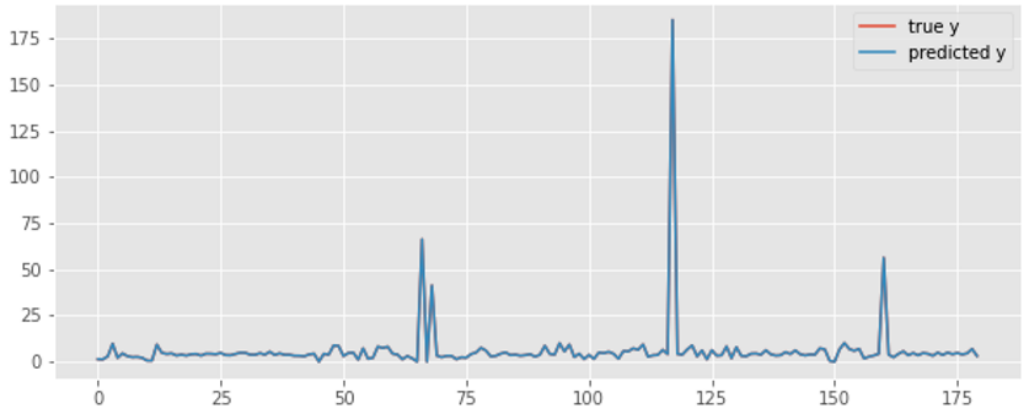


Figure 2. Weights(with extreme CT): old:76.1, dis:113.7, child:0.6 ,lowin:4.8, Leng:14.2, color:55.2, medical:18.7

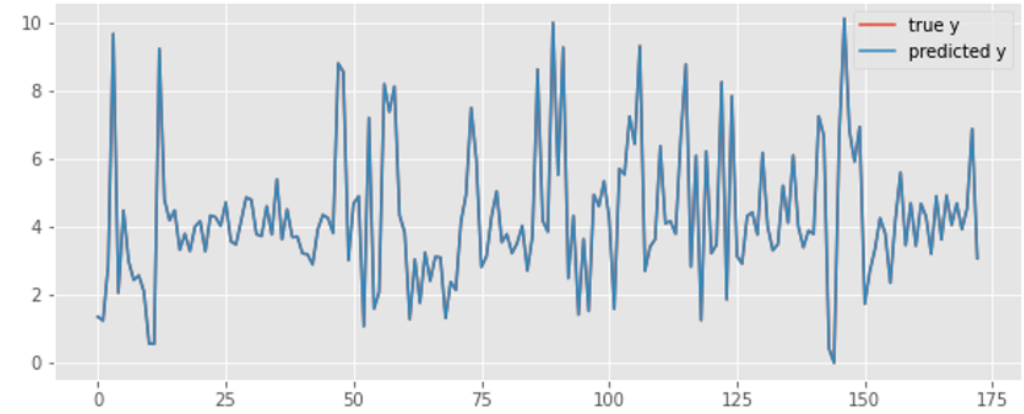


Figure 3. Weights(without extreme CT): old:6.1, dis:7.0, child:6.5, lowin:5.9, Leng:6.8, color:4.3, medical:9.8

Using those methods, we found that the medical group counts more in the request while the colored group counts less.

Analyze which cities submitted the most service requests

A huge part of this project is that we want to know which communities are most empowered based on 311 service requests. A good way to get the result is to analyze which insert geography submitted the most service requests. We decided to use CT ID instead of region to conduct our analysis because CT ID is more specific and precise if we want to get an accurate result. We broke down our analysis into three stages: CT with most service requests, CT with the greatest ratio of service requests to its population, and CT with over 500 service requests and the greatest ratio of service requests to its population. We decided to take the percentage and the number of total requests into consideration because it is more meaningful and unbiased to interpret the percentage when the data is larger. Below are the three graphs of the stages mentioned above. From the stage we can see that street cleaning, highway maintenance and abandoned vehicles are especially common reported in some census tracts. We've also found that census tracts with more colored people are commonly reported about sanitation, observing 34 of 36 census tracts with most coloured people most highly reported sanitation.

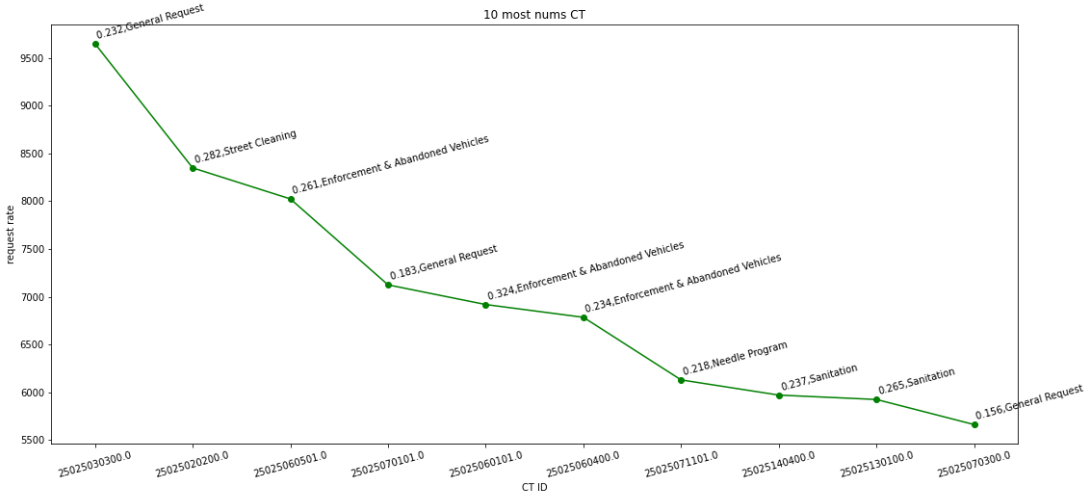


Figure 4. 10 CT IDs with most service requests

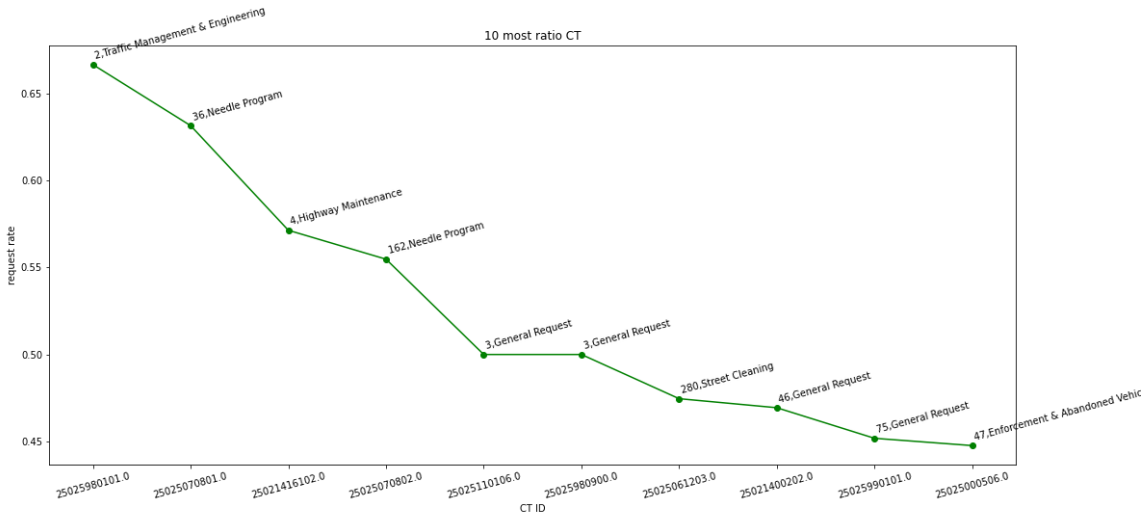


Figure 5. 10 CT IDs with greatest ratio of service requests to population

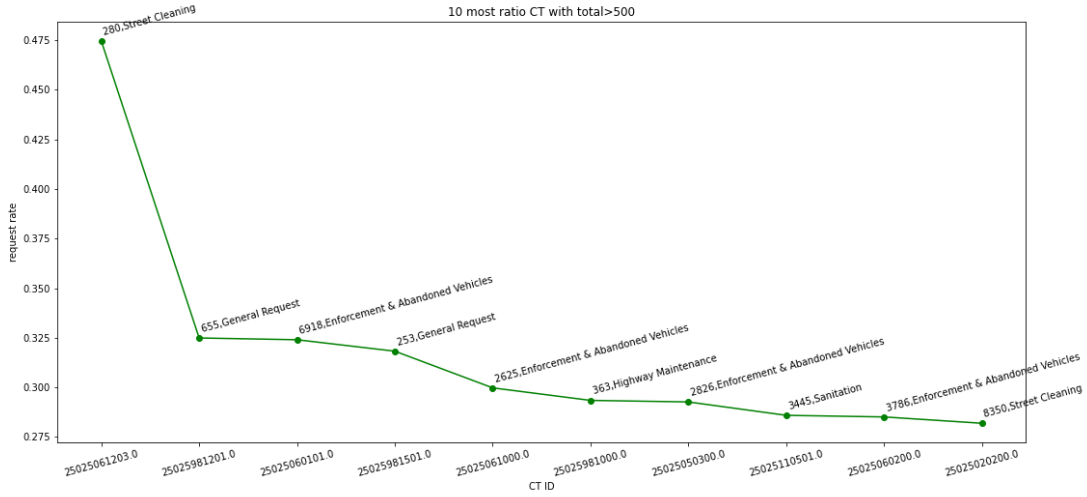


Figure 6. 10 CT IDs with more than 500 requests and with greatest ratio of service requests to population

Is the rate of closure for different types of service requests the same across [insert geography]?

No, not only do some types of requests have less closure rate than others, but also

certain kinds of census tracts would have less closure rate than others. We generated first-stage census tracts, types, and combinations of types and census tracts to discover the variation between types and CTs and distinguish those types or census tracts that has low closure rate and search for their common features.

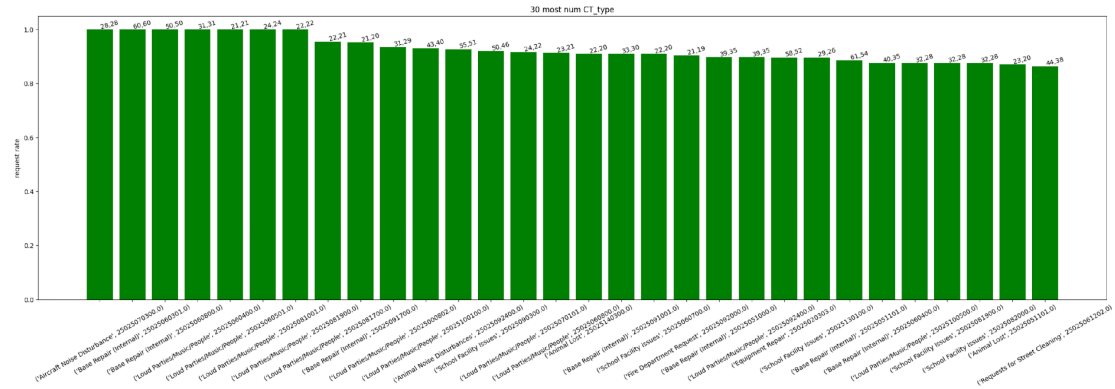


Figure 7. 30 CT ID_Type with more than 20 requests and with greatest ratio of unclosed cases

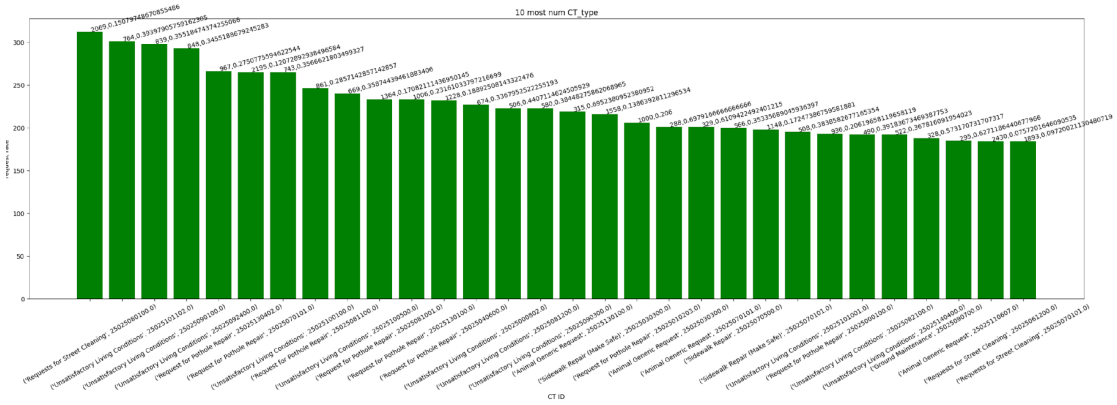
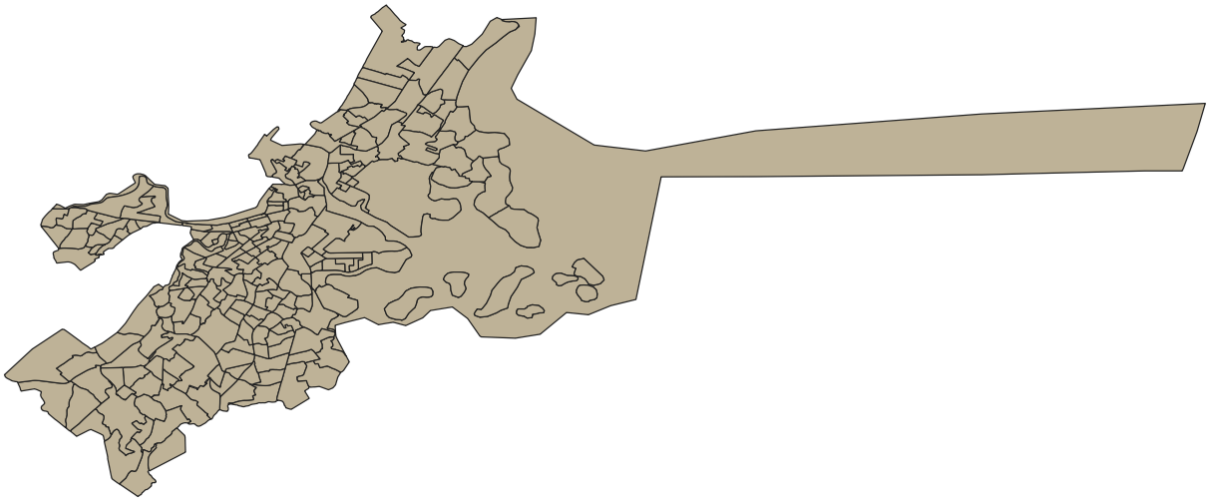


Figure 8. Most CT ID_Type with most numbers of unclosed requests

From the block graph, we found that for the census tracts with more cases to be received, the closure rate is even higher compared with those that do not receive that many requests. As we have found that with more people, there are likely to be more requests.

For further research on closure across CT, as we have distinguished the importance of geographic location, we used outside data from the America census center to get the

longitude and latitude of every CT. Therefore, we can get the relative position of CTs on a map. We create the heatmap the every point indicates a census tract and its scale shows the area of the CT. The color shows the unclosed rate. We did a similar map on different kinds of requests as the following figures show. We can see that regardless of which type of case, some certain CT's located between latitude (42.30,42.34) and longitude (-71.10,-71.05) always show a high unclosed rate.



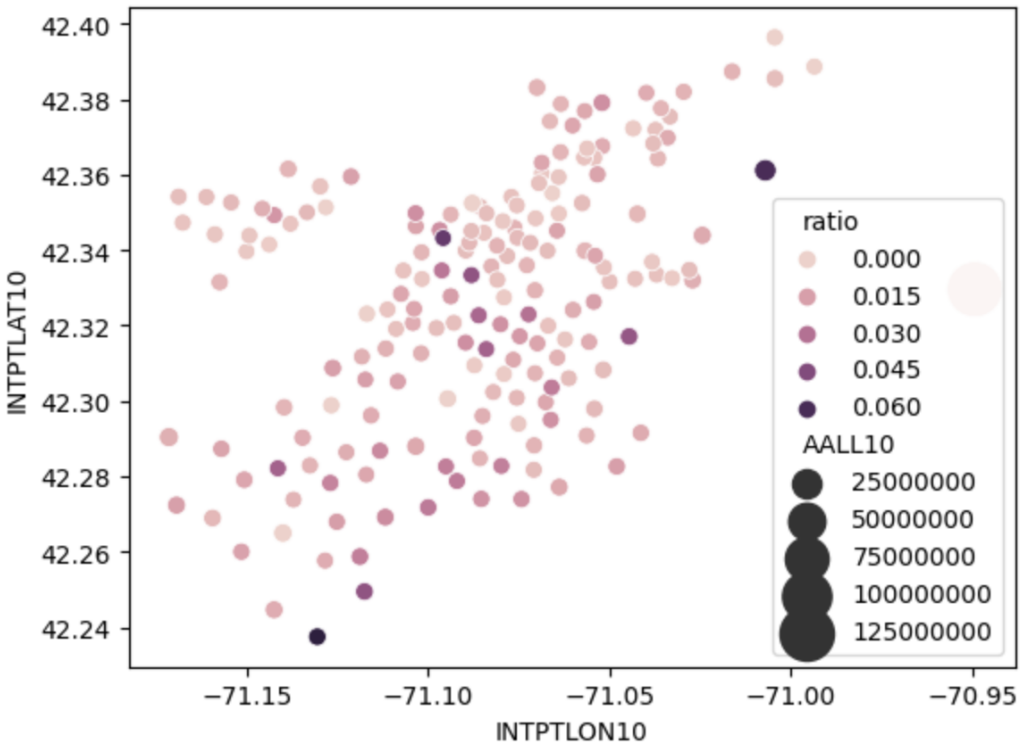


Figure 9. Parking Enforcement

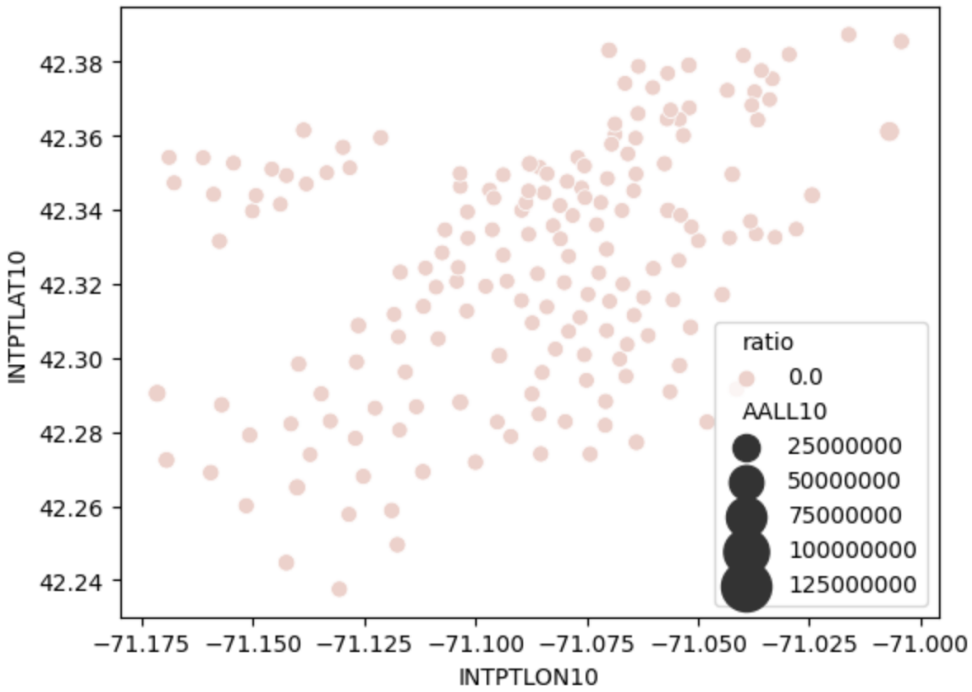


Figure 10. Scheduling a bulk item pickup

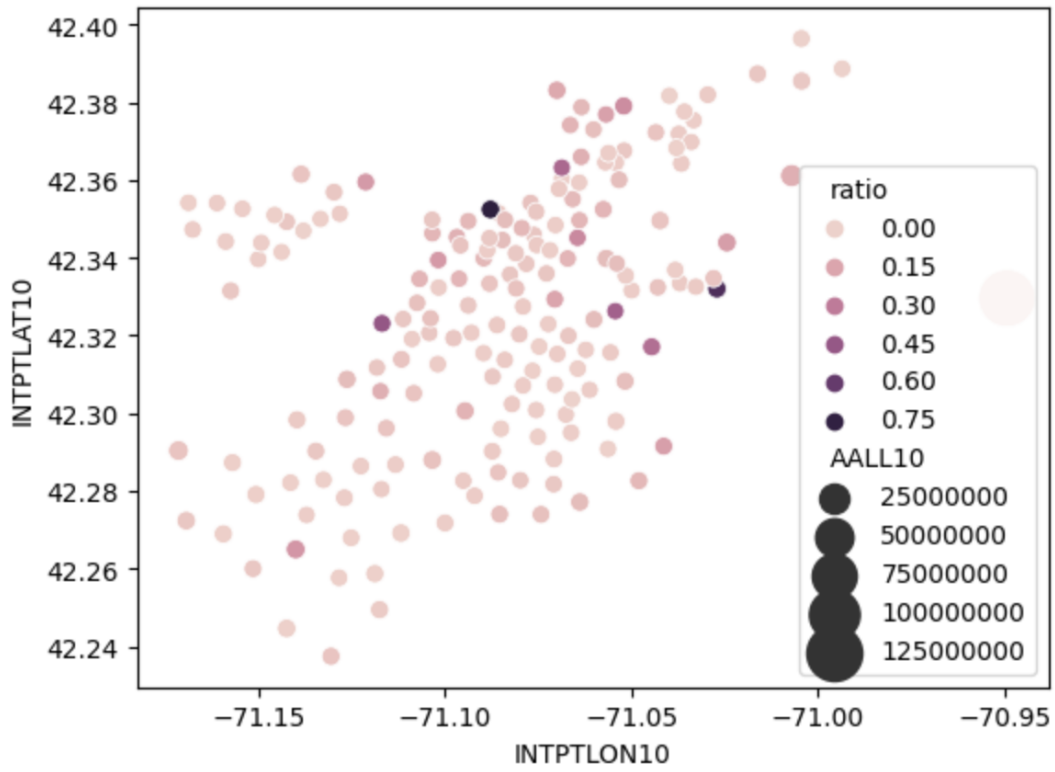


Figure 11. Requests for Street Cleaning

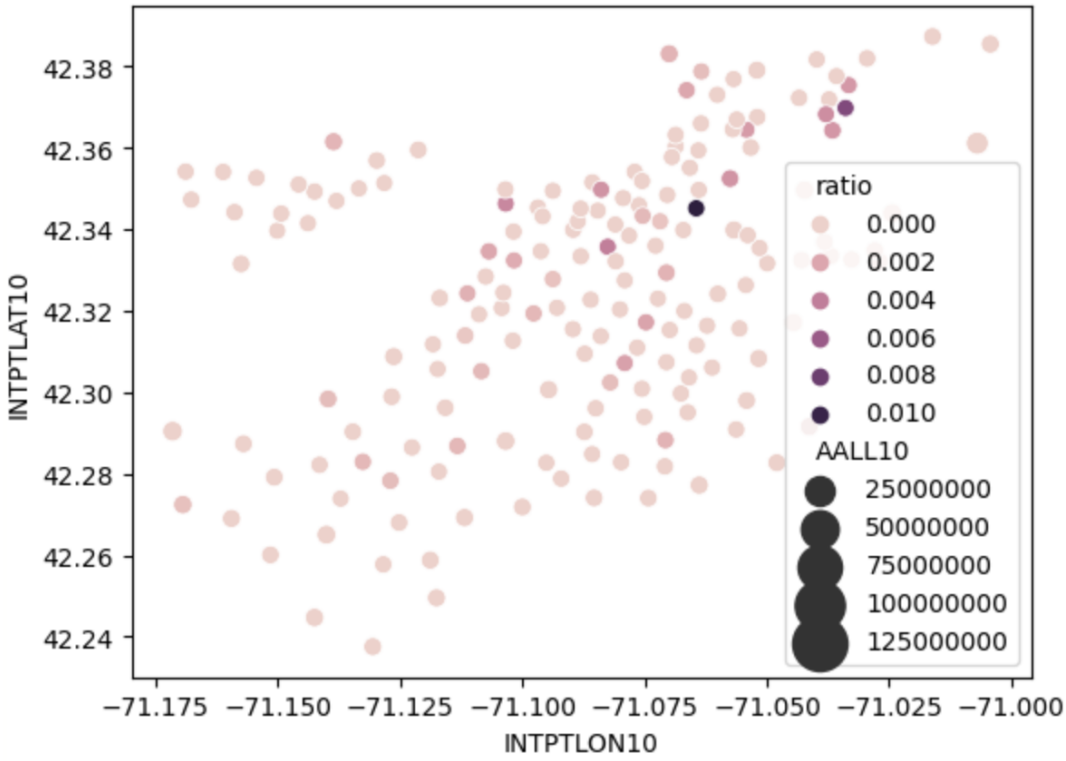


Figure 12. Missed Trash/Recycling/Yard Waste/Bulk Item

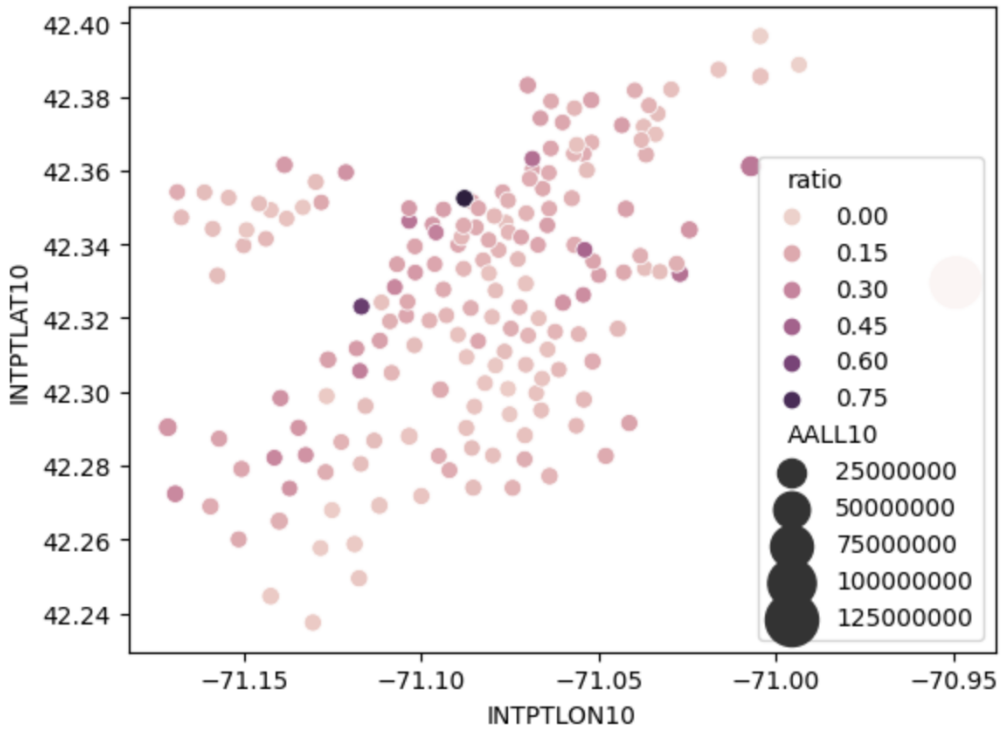


Figure 13. Requests for Pothole Repair

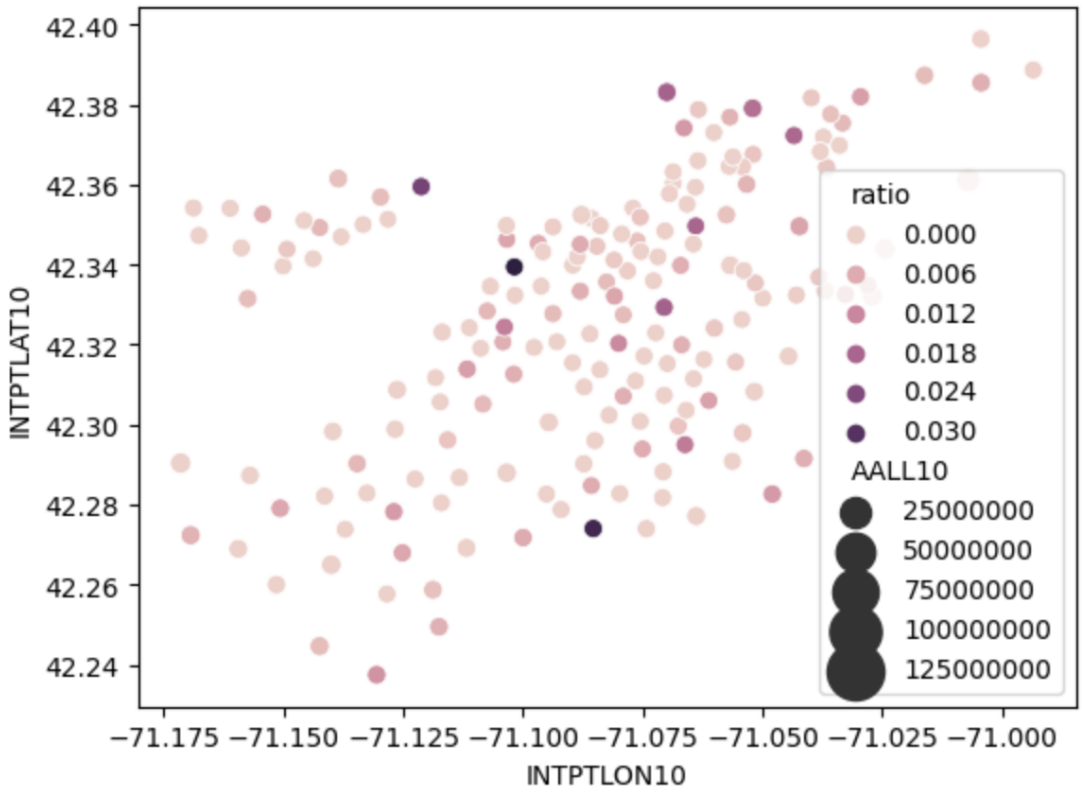


Figure 14. Improper Storage of Trash

Extension Analysis

Extension Project Goal and Analysis

So far we have stressed a lot on the relationship between social vulnerability and the number of service requests across different areas. To conduct a further analysis, we decided to explore what types of service do certain areas demand or lack. By doing this, we believe that the 311 service would be more comprehensive and satisfying.

Have the types of requests changed over the years for areas with high social vulnerability people?

In our base analysis, we figured out the areas with high social vulnerability through the calculation of social vulnerability index. Within those areas, we analyzed their service requests from the past ten years. In the graph below, we can see that the numbers of a couple types of requests such as General Requests and Parking Enforcement increased over the years while services such as scheduling a bulk item pick up went down a lot. This can be a possible indicator of whether some services require more resources or capitals.

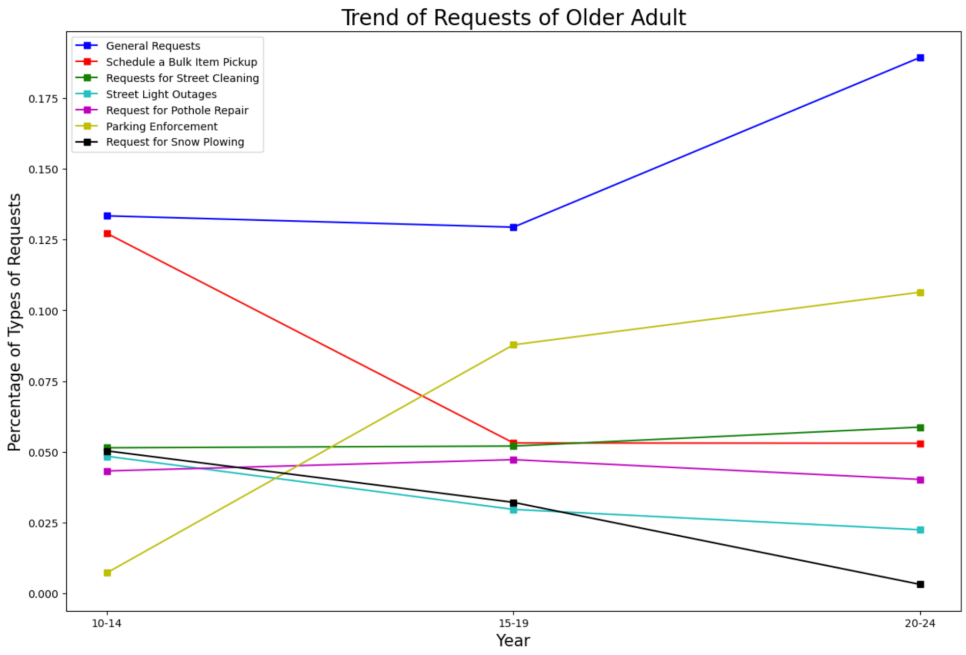


Figure 15. Trend of Service Requests of Older Adult

How are people reporting service requests?

By understanding how people are reporting requests, we can tell what forms are popular and efficient. In figure 6, we can tell that mobile requests are getting more and more popular. On the other hand, constituent calls are less common over the years. This result shows that more requests are coming from mobile calls and also implies that various types of service requests can be made through mobile. Although constituent calls are still an important format of reporting, more and more service requests are made by mobile calls due to its convenience. In addition, the City App has also been growing and has a lot of potential due to the technological advancements. Therefore, we think more emphasis should be placed on mobile calls and City App requests to make the system more well-rounded.

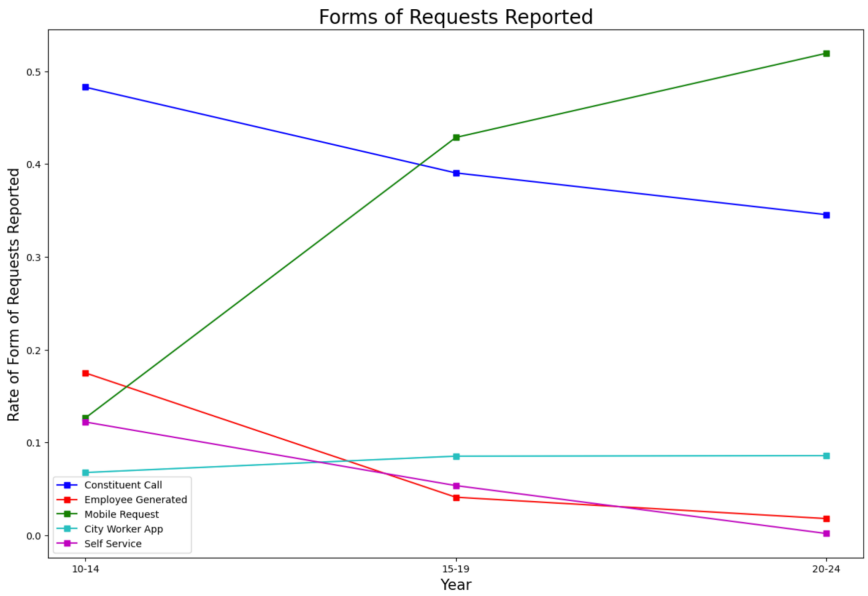


Figure 16. Forms of Requests over Years

Conclusion

We found that the closure rate of a census tract is highly correlated with the number of cases of the census tract. With different types of service requests, there are also different distribution patterns in different areas. It is useful to know that what types of service and what areas, based on the closure rates, require more attention and resources. We believe that our project provides insightful and meaningful information for the 311 requests organizers and our analysis can be used for a comprehensive service system.

Members' Contributions

Cheng-Ping Lin: data cleaning and preprocessing, SQL data sorting, analyze requests from different regions for base analysis, final report draft

Kyren Chen: data cleaning and preprocessing, SQL data sorting, analyze different types of requests from regions of base analysis, extension analysis

Weichen Jiang: data cleaning and preprocessing, python dataframe basic and advanced analysis and visualization on main analysis base on request, location, demographics ..etc