ANALYZING 311 RESPONSES

Members	Year	Email ID
Juan Guang	2023	jguang@bu.edu
Aishwarya Reddy Lachangar	2023	lareddy@bu.edu
Gaoyuan Han	2023	hgy@bu.edu
Anoohya Veerapaneni	2023	anuuv@bu.edu

Tables of Content:

2. Introduction

The goal of our project, "Analyzing 311 Responses," is to investigate the city's response to 311 service requests and assess whether they are being resolved equitably across different communities in Boston. This study is initiated by At-Large City Councilor Julia Mejia to gain a deeper understanding of which communities in Boston feel empowered to demand services and how the city addresses these demands from its residents. By examining 311 data, census data, and social vulnerability indexes, we will offer insights into the relationship between civic engagement, demographics, and city responsiveness. The broader impact of our analysis may shed light on potential inequalities in the allocation of resources and inform future policy decisions to promote a more equitable distribution of city services.

Overview

Our project, "Analyzing 311 Responses," aims to study the city of Boston's response to 311 service requests and examine whether these requests are being addressed equitably across various communities. By analyzing 311 data from 2011 to 2022, as well as appending relevant census data, we will gain insights into the demographics of communities that feel empowered to demand services and evaluate the city's responsiveness to these requests.

To achieve this, our analysis will cover several key aspects:

- Assessing the demographics and social vulnerability indexes of communities with different levels of service requests.
- Investigating the most common types of service requests for the entire city and various geographical areas.
- Analyzing the relationship between civic empowerment and factors like voter registration and voting history.
- Examining the city's responsiveness to 311 requests, including the rate and types of closures across different communities.

Through this comprehensive approach, we aim to shed light on any potential inequalities in the allocation of city services and provide valuable insights for future policy decisions that promote equitable resource distribution.

Project Goal

The project goal for "Analyzing 311 Responses" is to investigate the city of Boston's response to 311 service requests and assess whether they are being resolved equitably across different communities. By analyzing patterns between the demographics of communities, civic empowerment, and the city's responsiveness to these requests, we aim to identify potential disparities in the allocation of city services. This analysis will provide valuable insights to inform future policy decisions and promote a more equitable distribution of resources and services across all communities in Boston.

3. Base Analysis

Part 1:

Q. **Establish majority racial and ethnic composition of each 311-request based on block. group demographics. (ALL)**

We use two sources of data; one is the census data for Boston in 2020 and the other is the 311 data from 2022. The census data has populations in different races and ages for each precinct and each record in the 311 data also has precinct numbers associated with it. Therefore, it is very natural for us to connect these two datasets by the precinct numbers and using the precinct number to represent a neighborhood.

We first analyze the census data by precinct: We calculate the ratio of a certain population. over the total population in that precinct so for each 311 requests, we know which precinct it comes from and the racial composition of that precinct.

1. Conduct of analysis of requests:

We first focus on the on-time rate and the case-closed rate. We want to know if a certain racial composition of a neighborhood can affect the efficiency of the 311 service for that area. From the last step we can get the racial composition for each precinct, and we calculate the spearman correlation coefficients of the on-time ratio and the case-closed ratio of 311 service of all the precinct and the racial composition, and we have some findings. In terms of on-time, we found that the White and Hispanic people ratio in a community both have positive correlation coefficient with the on-time rate; while the Asian and Black people ratio both have negative correlation coefficient with the on-time rate. However, in terms of case-closed rate, only the Hispanic ratio has a negative correlation coefficient with the case-closed rate.

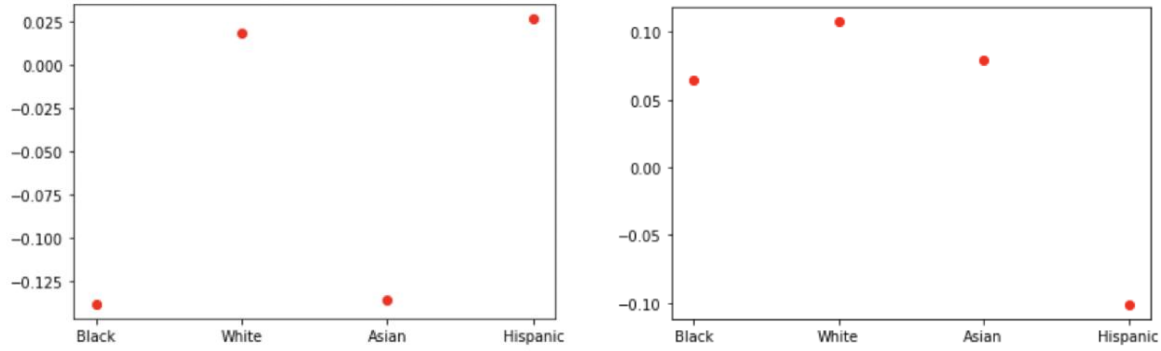


Figure 1. The left plot is the correlation coefficient of on-time rates versus respective racial proportion of communities: The right plot is the correlation coefficient of case-closed rates versus respective racial proportion of communities.

Then we try to figure out what kinds of cases a certain race of people tends to report to 311. The way we choose to analyze the problem is we focus on one race at a time: first we classify the precincts by population of a certain race into 3 clusters using k-means, each of

which represents low, mid and high proportion of that race. Then we plot the number of different cases for the 3 clusters.

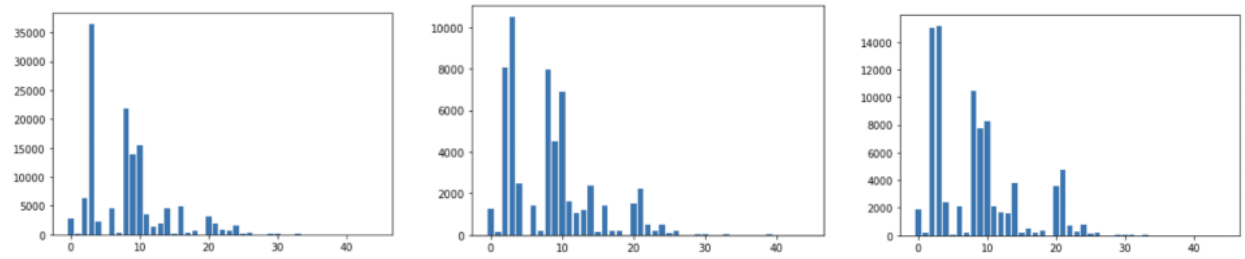


Figure 2. The left plot is the number of different cases for low white people population communities; The middle plot is the number of different cases for middle white people population communities; The right plot is the number of different cases for high white people population communities.

We first focus on the White population (as in Figure 2), the shape of the number of different types of cases seems to be similar, but we notice that the third bar in the left plot is lower than the ones in others. The third type corresponds to the 'Sanitation'. Therefore, it seems like that for communities with lower White population, cases about Sanitation is reported less. The reason could be there are less phenomena about sanitation in those communities or people of this race report more often, or something that is more complicated.
We perform the same procedure with respect to Black, Asian, and Hispanic race. For Black race, we notice a similar pattern, however, for Asian and Hispanic, we notice an opposite pattern community with a high population of Asian and Hispanic report much fewer cases regarding sanitation as shown in Figure 3. The reason also could be complicated.

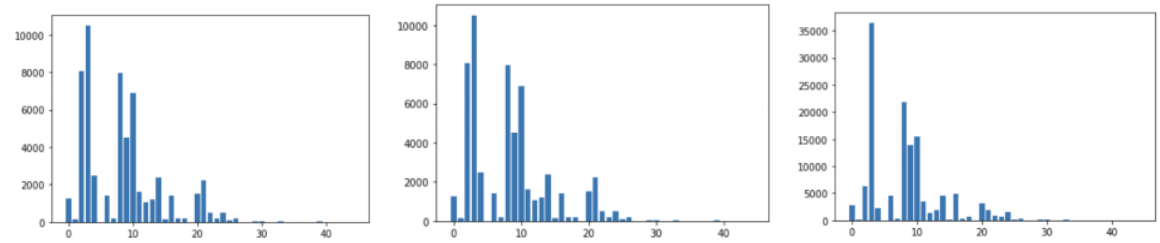


Figure 3. The left plot is the number of different cases for low hispanic people population communities; The middle plot is the number of different cases for middle hispanic people population communities; The right plot is the number of different cases for high hispanic people population communities.

Part 2:

Q. Which communities are most empowered based on 311 service requests? (All)
In this analysis, we aim to determine which communities in the city are most empowered based on the number and types of 311 service requests they submit. By examining the distribution of service requests across different neighborhoods and their corresponding demographics, we can gain insights into the communities that feel more confident in demanding services from the city. Understanding these patterns will help inform

policymakers and city officials on where resources may be best allocated to address the needs of all residents and promote civic engagement across the city.

Q. What is the pattern of demographics and service requests: looking at census block groups and social vulnerability indexes?

i. Different types of requests (Guang & Han):

Communities with higher social vulnerability index scores (higher vulnerability) generally have lower numbers of service requests, indicating that less empowered communities might be less likely to utilize the 311 system. There is a negative correlation between the social vulnerability index and the number of service requests. This means that as the social vulnerability index increases (more vulnerable communities), the number of service requests tends to decrease. Here in the below table we took, top 5 reasons(Sanitation, Street Cleaning, Enforcement & Abandoned vehicles, Highway Maintenance and Code Enforcement) as these seemingly have most of the high cases when compared with other reasons. We analyzed each of these reason cases with respect to the census block group.

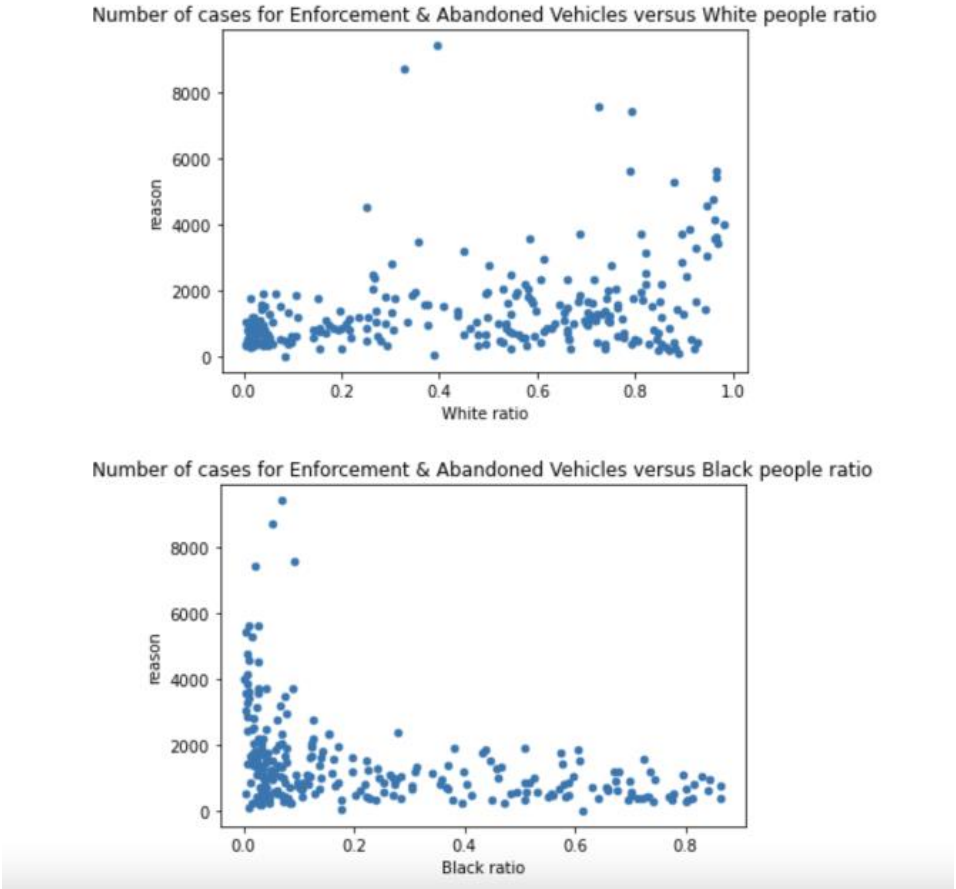


Fig: example of the scatter plot which helps us to under the coefficients in the below table.

With the help of the scatter plot 1, The purpose of the scatter plot is to identify trends and potential correlations between these two variables. The overall pattern suggests a strong positive correlation, meaning that as the values of White Ratio increase, the values of Reasons tend to increase as well. The distribution of data points appears to be linear, with a tight clustering around a central line. There are no significant outliers present in this scatter plot, indicating a consistent relationship between the two variables.

For scatter plot 2, The purpose of the scatter plot is to investigate the trends and potential correlations between these variables. The overall pattern does not display any clear correlation, suggesting that there is no significant relationship between Black ratio and Reason. The data points are widely dispersed, and there is no discernible trend or shape to their distribution. There are no significant outliers in this scatter plot, as the data points are already broadly distributed and do not follow any specific pattern.

	Sanitation	Street Cleaning	Enforcement & Abandoned Vehicles	Highway Maintenance	Code Enforcement
White ratio	-0.0246	0.0418	**0.3288**	0.2671	0.1957
Black ratio	0.1845	-0.1125	**-0.4417**	**-0.3091**	**-0.3036**
Asian ratio	-0.2873	-0.0183	0.1907	0.1564	0.1283
Hispanic ratio	-0.1448	0.0274	-0.1895	-0.0486	-0.1284
White ratio (over 18)	-0.1210	0.1395	**0.3806**	**0.3729**	**0.2747**
Black ratio (over 18)	0.1829	-0.0143	**-0.3715**	-0.1926	**-0.2221**
Asian ratio (over 18)	**-0.2851**	0.0863	**0.2589**	**0.2529**	**0.2073**
Hispanic ratio (over 18)	-0.1631	0.1460	-0.0725	0.1506	-0.0120

Fig: Shows the coefficients correlation between races ratio with the top 5 different types of cases

- Each number(coefficient) represents a story for the communities in a dimension.
- For example, 0.3288 shows a positive correlation, it means that if there are more white people in a community, then there tends to be more cases for Enforcement and Abandoned Vehicles reported.
- Similarly, -0.4417 shows a significant negative correlation, meaning that if there are more black people in a community, then there tends to be less cases for Enforcement and Abandoned Vehicles reported.

The next table shows the coefficient correlation between different reasons and the different dimensions in the Social Vulnerability Index (SVI)

	Sanitation	Street Cleaning	Enforcement & Abandoned Vehicles	Highway Maintenance	Code Enforcement
OlderAdult ratio	**0.2991**	**0.1937**	**0.1179**	**0.2671**	**0.1409**
TotChild ratio	**0.4116**	**0.1648**	0.0174	0.0515	-0.0143
LEP ratio	-0.0929	-0.0514	0.0276	-0.0715	-0.0401
Low_to_No ratio	**-0.1268**	-0.0316	0.0236	-0.0945	0.0026
MedIllnes ratio	0.1105	**0.2088**	**0.2535**	**0.3302**	**0.3035**
Disability ratio	**0.2477**	0.0874	0.0476	0.0601	-0.0187

Fig: The figure shows the top 5 different reason cases coefficient correlation with respect to the different communities (SVI) in boston city.

- We can see that for Older adult ratio and Children ratio, the coefficients are all significantly positive, this means that if there are lots of children and elder people in a community, then there tends to be more cases about Sanitation and Street Cleaning reported.
- Speculation: it is reasonable since if there are lots of children and elder people, then people who take care of them probably want them to live in a clean environment, which leads to more attention about Sanitation and Street Cleaning

ii. **All requests (Aishwarya):** The total number of service requests across all neighborhoods in the dataset is 40,183. Allston/Brighton, Jamaica Plain, and Dorchester are the top 3 neighborhoods with the most service requests. These communities might be more likely to engage with the 311 system and could be considered more empowered based on their high number of service requests.

Even in this section we took the same steps as we did in the above for different types of reasons.

Different races	Correlation ratio between 311 case and different races	Correlation ratio between 311 case and different races over 18 years
White	0.2469	0.2432
Asian	0.0415	0.0411
Black	-0.2938	-0.2923
Hispanic	-0.1335	-0.1299

Fig: The above table shows the coefficient correlation between the 311 cases with age over 18years and all together with respect to different races.

The 0.2469 & 0.2432 shows Positive correlation that means, as one variable increases the second variable increases too. This means that if the white people population increases than the number of total cases increases too in the boston city. Similarly, -negative correlation -0.2938 & -0.2923 means that they are inversely related, one increases and the other decreases. That means with the data we have that even if the Black population increases the total number of cases 311 cases have negative correlation, that number of cases decreased as the population increased. This might say many different things but, the above point and table just showcases what we got from the data we received.

Similarly,

Communities	Correlation ration between 311 cases and
Older adult	0.1465
Disability	-0.0462
People of colour	-0.1179
Limited Eng proficiency	-0.0902
Medical illness	0.2921
Low to no income	-0.1494
Child	0.0679

fig: Correlation coefficient correlation between different svi communities with total number of 311 cases

From the above table we could say that medical illness has a positive correlation whereas the low to no income has the negative correlation of -0.1494.

Q. How does the city respond to its residents based on the ratio of resolved to unresolved 311 requests? (Guang)

In this analysis, we seek to understand how the city responds to its residents by evaluating the ratio of resolved to unresolved 311 service requests. By investigating the response rates and types of requests that are resolved or remain unresolved, we can determine the efficiency and effectiveness of the city's response to various neighborhoods and demographics.

But here we hit the data limitation because, to see which requests are closed, we should be using the requests which have a closed date and a closure reason. The ones which don't have a closure reason or say "case closed/case noted" will be a case in which we're not sure whether the case is closed or not. Please use the requests that are closed, and you can say there are certain requests which we're not sure of, which is a data limitation.

Data trimming:

- Discard the data that:
- Don’t have a closed date.
- Don’t have a close reason.
- “Case closed/case noted” for close reasons.

With the above cleaning we have trimmed our dataset with 10,000 cases. But with this we were able to move forward with our project without any error analysis.

Results:

i. With Census:

```
corrs
```

	All	Enforcement & Abandoned Vehicles	Sanitation	Code Enforcement	Street Cleaning	Highway Maintenance
White	0.230045	0.456998	0.448858	-0.073420	-0.189162	-0.362412
Black	-0.274822	-0.594802	-0.444462	0.012861	0.130318	0.265128
Asian	-0.027337	0.454765	0.283033	-0.107233	-0.169419	-0.289327
Hispanic	-0.165900	-0.323966	-0.285845	0.045371	0.096809	0.363985

Similarly, how we found the coefficient correlation for total number of 311 requests with the different race’s communities, here in this we try to find the coefficient correlation between closed cases with the different communities plus also top five different reasons for cases in Boston. Here by observing the above table we can see that white race population tends to have positive correlation, in that most of it is contributed by Enforcement & Abandoned vehicles which is around 0.456, whereas Highway Maintenance has a negative correlation with -0.3624.

The Black population tends to have a negative correlation of -0.2748 respectively, but here we can see that Enforcement & Abandoned vehicles have about -0.5948 correlation whereas Highway Maintenance has a positive correlation of 0.26512.

When we compare the table overall, we can see that White and Black populations are inversely proportioned.

ii. With SVI:

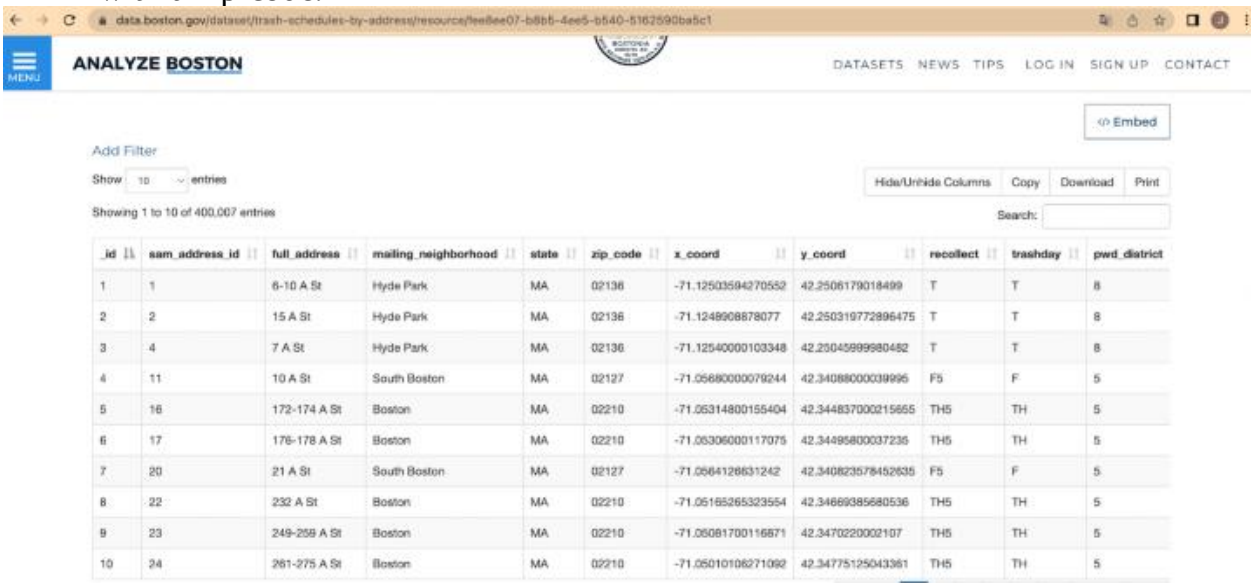
	All	Enforcement & Abandoned Vehicles	Street Cleaning	Sanitation	Highway Maintenance	Code Enforcement
OlderAdult ratio	0.168130	-0.068118	0.058348	-0.013723	-0.016511	-0.002410
TotChild ratio	0.007721	-0.472647	0.227310	-0.295865	0.296077	0.028548
LEP ratio	-0.120462	-0.138799	0.077234	-0.272038	0.143849	0.020858
Low_to_No ratio	-0.088924	-0.092859	0.114258	-0.313566	0.165102	0.070150
MedIllnes ratio	0.292094	0.317999	-0.086331	0.219265	-0.230224	-0.005921
Disability ratio	-0.075978	-0.324719	0.112694	-0.283035	0.227647	0.019307

In the same fashion, we applied it to SVI and different types of reasons. Medical illness ratio has the positive correlation, where Enforcement & abandoned vehicles have positive correlation, but Highway Maintenance has negative correlation. And Low to no income population has a negative correlation.

4. Exploratory Analysis**(Guang)**

We have already analyzed the 311 data with respect to the composition of communities, including the racial and age composition, and the dimensions included in SVI. From our analysis, especially with the 311 data, I think we have some significant and helpful findings that tell us about communities and their needs in Boston. Then it is very natural for us to think about the government's awareness and response to the 311 service and the composition of communities.

We notice that Sanitation and Street Cleaning are very frequent cases from 311 data, so we want to see the performance of the corresponding service. We have data about the current trash collecting schedule. As shown in the figure, the 'recollect' and 'trashday' columns indicate the frequency of trash recollection and trash collection for a location. The values are interpreted as below: M means every Monday, T means Every Tuesday, W means every Wednesday, TH means every Thursday, and F5 or TH5 means every fifth Friday or every fifth Thursday. With this data, we can calculate the frequency of trash collection every week for every location marked with a zip code.



_id	sam_address_id	full_address	mailing_neighborhood	state	zip_code	x_coord	y_coord	recollect	trashday	pwd_district
1	1	6-10 A St	Hyde Park	MA	02136	-71.12503594270552	42.2506179018499	T	T	8
2	2	15 A St	Hyde Park	MA	02136	-71.1248908878077	42.250319772896475	T	T	8
3	4	7 A St	Hyde Park	MA	02136	-71.12540000103348	42.25045999980482	T	T	8
4	11	10 A St	South Boston	MA	02127	-71.05680000079244	42.34088000039996	F5	F	5
5	16	172-174 A St	Boston	MA	02210	-71.05314800155404	42.344837000215655	TH5	TH	5
6	17	176-178 A St	Boston	MA	02210	-71.05306000117075	42.34495800037235	TH5	TH	5
7	20	21 A St	South Boston	MA	02127	-71.0564128631242	42.340823578452635	F5	F	5
8	22	232 A St	Boston	MA	02210	-71.05165265323554	42.34669385680536	TH5	TH	5
9	23	249-259 A St	Boston	MA	02210	-71.05081700116871	42.3470220002107	TH5	TH	5
10	24	261-275 A St	Boston	MA	02210	-71.05010106271092	42.34775125043361	TH5	TH	5

Figure. An example of trash schedule data

We calculate the close rate and case number for Street cleaning and sanitation. We interpret these two numbers as a community's need of the service for Street cleaning and sanitation. We calculate the frequency of trash day and recollect for every community by zip code. We interpret these two numbers as the government's devotion to cleanliness for every community. Then we plot the close rate and case number versus the frequency of trash day and recollect for every community.

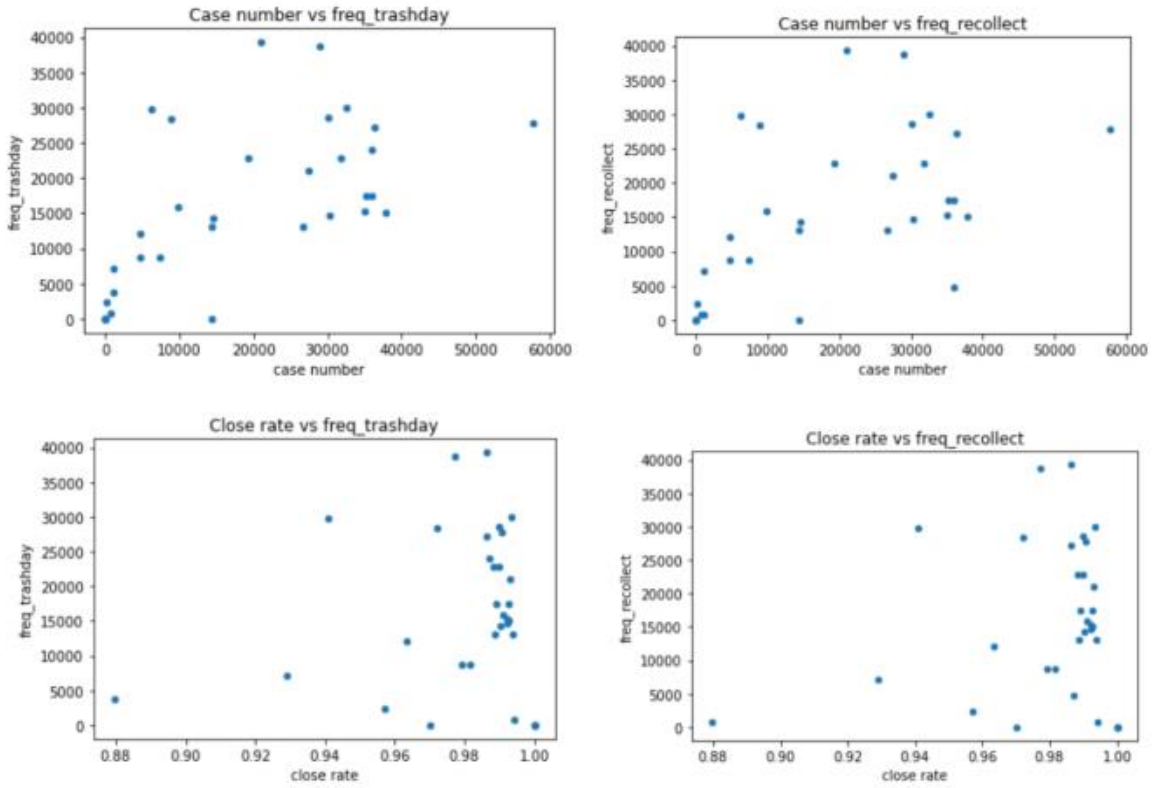


Figure. Scatters for Close rate and case number versus frequency of trash day and trash recollection.

As we can see in the four scatter plots, the trash collecting service performs well with respect to the total case number of Sanitation and Street Cleaning for every community: Communities that have only a few cases about Sanitation and Street Cleaning reported get less trash collecting service, while communities that have a lot of cases about Sanitation and Street Cleaning reported get more trash collecting service. However, we found that with respect to close rate for Sanitation and Street Cleaning, there are two outlier communities where the close rate for Sanitation and Street Cleaning is low but the frequency of trash service is still not high enough. These two outliers are communities with zip codes 02210 and 02110.

5. Conclusion

Throughout the project, we had established a majority racial and ethnic composition of 311 requests based on block group demographics, which comes from the census data and social vulnerability data. From this preliminary work, we learned to analyze a community by looking at its demographic composition and intuitively answer the question: How does 311 service benefit communities in Boston? We answer this question in various aspects. We

think about how 311 service in general empowers communities by looking at numbers of all the cases reported from each community and the close rate for cases from each community, then we calculate their correlation with the demographic composition. We also choose the 5 most frequent cases reported to conduct similar research. We found that for some types of cases, its performance, described by number of cases reported and close rate, display vastly differently among communities of different demographic composition. For example, we discovered a positive correlation with White people proportion and the close rate for Sanitation cases, while a strong negative correlation is shown with Black people proportion and the close rate for Sanitation cases. All the discoveries were contained in the correlation tables we constructed in this paper.

For the extension analysis, we want to evaluate the government's additional performance given the performance of 311 services for Boston communities. We notice that Sanitation and Street Cleaning are very frequent cases from 311 data, so we want to see the performance of the corresponding service. We use a trash collecting schedule to calculate the frequency of trash collection for each community and compare it with the case number and close rate for Sanitation and Street Cleaning cases. We evaluate the trash collection service to be efficient since, other than a few outliers, the communities with a lot of Sanitation and Street Cleaning cases reported tend to have more frequent trash collecting service and the communities with low close rate for Sanitation and Street Cleaning also tends to have more frequent trash collecting service.

6. Work Division

Team members	Work
Juan Guang	Data cleaning and Processing, Base analysis Demographics and service requests(different types of requests), Exploratory analysis, Presentation & draft report.
Aishwarya Reddy Lachangar	Initial cleaning and processing, data collection, Base analysis (communities empowered based on data), comparing with all requests, Presentation, and Draft Report.
Gaoyuan Han	Base analysis, Demographics, and service requests (different types of requests)
Anoohya Veerapaneni	Base analysis (SVI and Census), visuals representation.