Analyzing 311 Responses

Ivan Niktitovic, Stephen Wong, Haoxuan Sun
in@bu.edu, swluquez@bu.edu, hs4379@bu.edu

April 14, 2023

Introduction (Worked by Stephen)

With the collaboration of the Boston City Council and Councilor Mejía, the "Analyzing 311 Calls" project deals with helping understand relationships between 311 callers and their correlation with the events that happen in the community. For this project, we are studying more specifically how people more likely to do 311 calls are correlated with leaders in the community who are more likely to have active voices in the community and feel empowerment. We also want to see the effectiveness of how these 311 calls are answered, whether they are answered in a responsible time frame or if they get to be solved at all. We use data from the main areas in Boston to check where there is an increase in calls, where people are most likely to vote and other factors that may affect the community. Our goal for this project is to help Councilor Mejía by providing information that can be useful to combat problems in the community by allocating resources towards areas with more need for response and also provide a correlation between empowerment, voting status and making 311 calls.

Base Analysis (Worked by Stephen, Ivan)

For the analysis, we first worked on getting all the data together. We used the datasets for 311 calls from 2011 to 2022 to conduct our analysis. We also pulled information from census data, vulnerability status information, and voting turnout to answer some of our other questions. This information wasn't always one to one with the other so a good amount of cleaning was done.

Our first question was:

1. Analyze which neighborhood submitted the most service requests?

From table 1 we can observe that the data for most submitted service calls is very much popular in a subset area of Rosindale. The data we pulled classified different populations within an area as different hence there is repetition on our top 10 most calls per population and doesn't refer to the whole area needing help but more of a subset. This information paired with the cluster from Table 2, shows that these areas in the top 10, over the past 10 years or so have been needing the most response to 311 calls.

For our second question we studied:

2. What is the relationship between voting status (i.e. registered voters and and civic empowerment? What is the average voting turnout for each district over past XX years?

For this question we paired different elections with social vulnerability index and the voting turnout for every year starting at the 2002 election to see the vulnerability index of the areas who voted during each election. In Table 3 we have an example of the turnour for 2002 in which we can observe that in most areas there was around 40 to 80 percent of each area voting and we see that it was mostly from people on the 0.2 to 0.6 on the vulnerability index. This shows that people on either end of the vulnerability index scale are either less likely to vote or that most people in an area are found somewhere in the middle of the scale. In Table 6 we did a linear regression in which we can see that the higher the population in the Social Vulnerability Index the less likely the turnout.

On table 4, we have an example of 5 areas in Boston and vulnerability index compared to the voting turnout on each election as well as the amount of 311 calls done per area. We can check here again our study from the previous question in which Social Vulnerability is linked to calls and we get to see what percentage of the votes this subset area did for the general area.

For our third question we studied:

3. What is the relationship between voting status (i.e. registered voters and and city responsiveness in terms of closing 311 requests? What is the average voting turnout for each district over past XX years?

For this question, similarly to the previous one, we take voting status again but compare it to its relationship with the closing of the 311 requests. As we can see in Table 6, we see there isn't really a big correlation between the two. We see most calls take around 20-25 days to get answered and also that election turnout ranges more between 30 to 50 percent.

Extension Analysis - Property Violations (Worked by Haoxuan and Ivan)

For our extension part, we raised the idea of analyzing the relationship between building/property violations and 311 requests using the Boston government dataset. With this idea, we could find areas where building violations are more frequent, therefore being able to set more specific resources towards those areas. This would be more specialized resources rather than easier requests such as graffiti and trash removal.

There are two reasons why the team chose this extension idea. Firstly, building and property violations impact public safety and quality of life. By identifying contributing factors, the team hopes to aid in developing effective prevention and addressing strategies. Also, 311 requests provide insight into residents' concerns and priorities. Analyzing their relationship to building violations can inform policy decisions and improve Bostonians' lives.

Two questions have been analyzed by now:

1. Is there a pattern between neighborhoods with more property violations and 311 requests made?

   According to our research, we found there to be one. It is shown on the heatmaps in Table 7 where we can observe the number of property violation records since 2009 (left) and the 311 calls from 2010-2023 (right) on the Boston neighborhood map . It is obvious that the neighborhoods with more property violations tend to have more 311 calls, which is reasonable.

2. What areas have the most property violations?

   The bar chart in Table 9 and pie chart in Table 8 indicate that Dorchester has the highest number of violations compared to other neighborhoods, accounting for approximately 28% of all violations. East Boston and Roxbury follow as the second and third highest neighborhoods in terms of violations. Together, these three neighborhoods account for almost half of all violations.

   By looking at the violations that happened in only Dorchester and group by zip code (Table 10), we can find out that the area with zip code “02124” had most violations and accounts for nearly a third of all property violations in Dorchester and in Table 11 we can observe how they are divided.

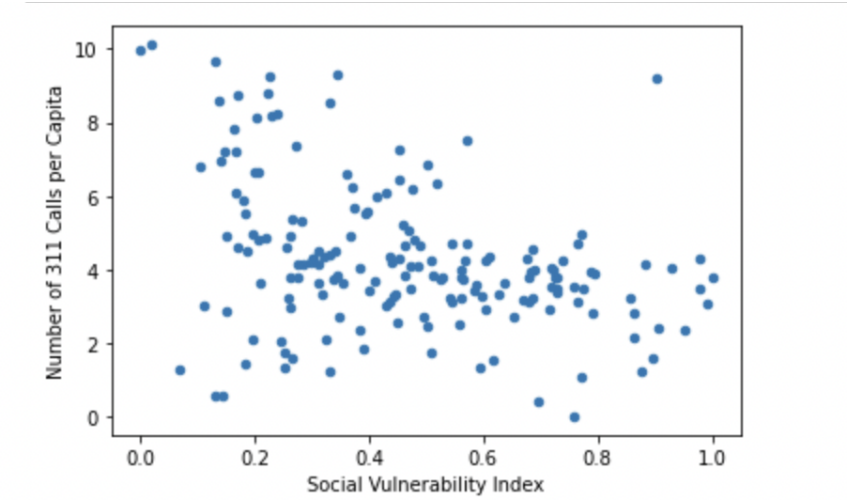
Appendix (Worked by Stephen, Ivan, Haoxuan)

Table 1

	CT_ID_10	NUM_CALLS	POP100_RE	Name	SV_IDX	CALLS_PER_POP
169	25025980700	1110	6	Roslindale	0.710523	185.000000
175	25025981501	795	12	Back Bay	0.072193	66.250000
170	25025981000	1237	22	Roslindale	0.187461	56.227273
179	25025981800	3383	82	Jamaica Plain	0.079128	41.256098
176	25025981502	91	9	East Boston	0.018031	10.111111
173	25025981202	2067	207	South Boston Waterfront	0.000000	9.985507
29	25025010702	23186	2400	Back Bay	0.130397	9.660833
72	25025070300	36366	3909	Bay Village	0.345587	9.303147
46	25025040600	22631	2444	Charlestown	0.227252	9.259820
168	25025980300	3117	338	Jamaica Plain	0.902063	9.221893

Table showing areas with the most 311 requests

Table 2



Scatter plot of the social vulnerability index given the number of 311 calls per group

Table 3

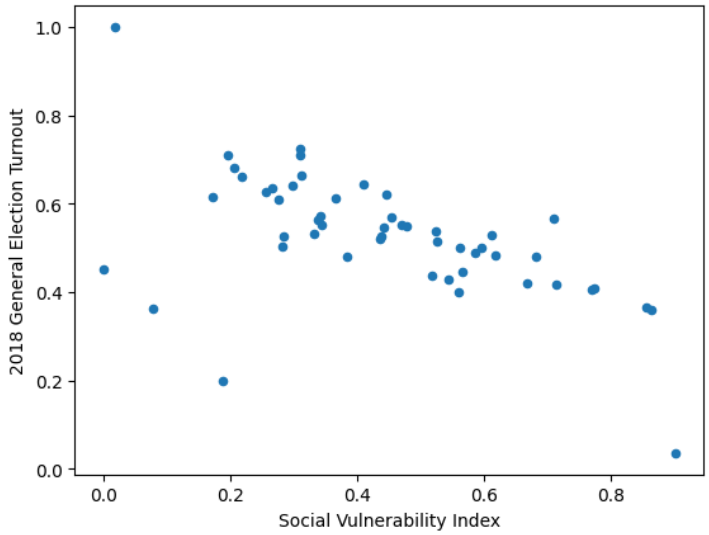
```
svidx_voter_turnout = calls_svidx_df.join(turnout_per_year_by_GEOID10.set_index('GEOID10'), on='CT_ID_10')

subset_svidx_voter_turnout = svidx_voter_turnout.dropna()

subset_svidx_voter_turnout.head()
```

[266] ✓ 0.0s Python

	CT_ID_10	NUM_CALLS	POP100_RE	Name	SV_IDX	CALLS_PER_POP	General_2022	General_2018	General_2014	General_2010	General_2006	General_2
121	25025100100	15519	5510	Roxbury	0.863899	2.816515	0.0	0.358926	0.273302	0.248341	0.201580	0.132
122	25025100200	9033	2787	Mattapan	0.559635	3.241119	0.0	0.398851	0.285057	0.279310	0.249425	0.168
123	25025100300	11487	3303	Mattapan	0.774576	3.477748	0.0	0.407569	0.307525	0.295781	0.264898	0.180
124	25025100400	17362	4865	Mattapan	0.587071	3.568756	0.0	0.488344	0.353012	0.338783	0.296397	0.208
125	25025100500	28227	5989	Dorchester	0.545204	4.713141	0.0	0.430067	0.290435	0.255527	0.219688	0.159

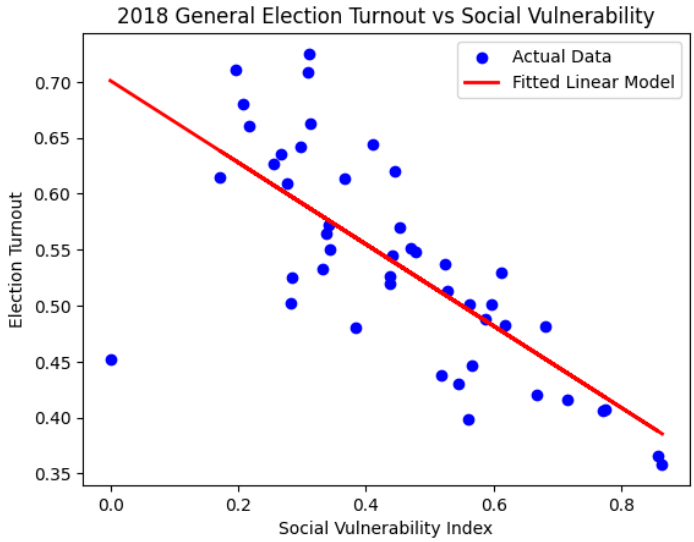
Table showing information about voter participation per area

Table 4



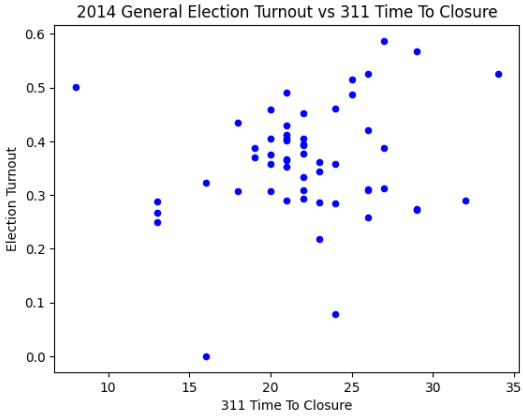
Sample of the 2018 election turnout related to the social vulnerability index

Table 5



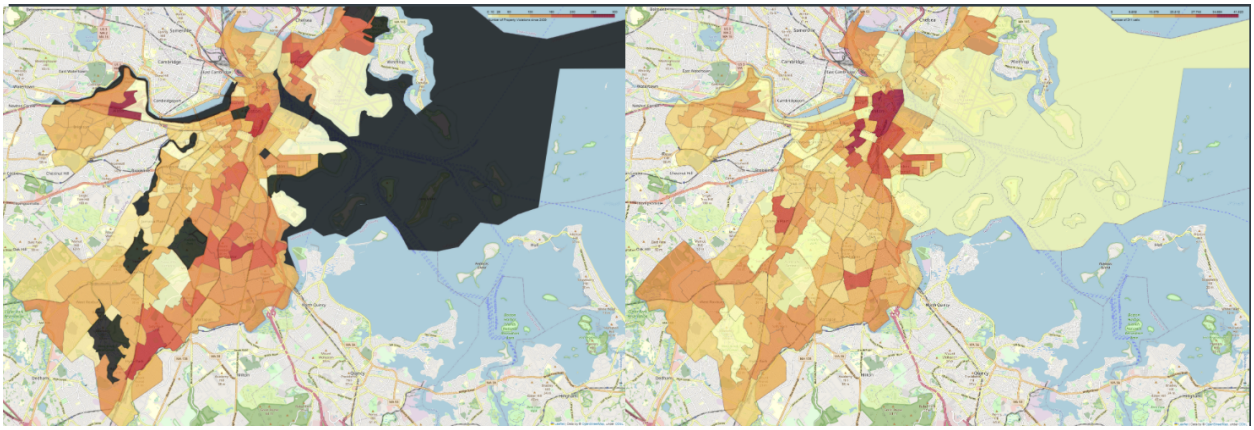
Linear regression of turnout and social vulnerability index

Table 6



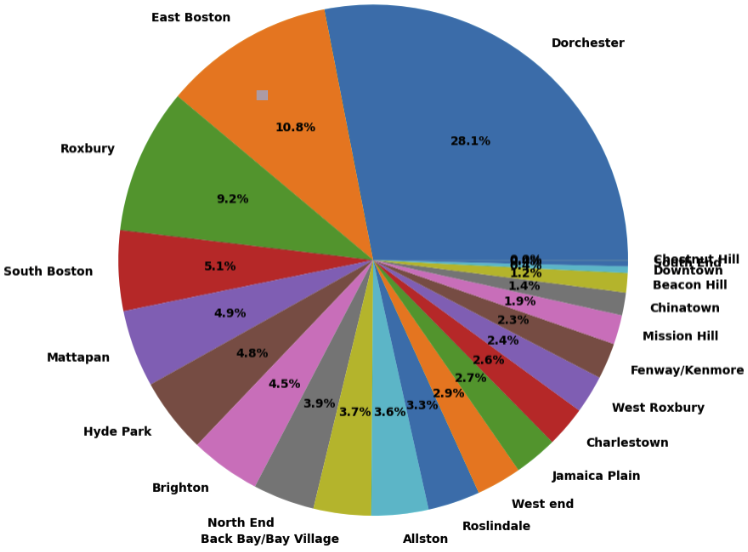
Scatter plot of 311 response times and election turnout

Table 7



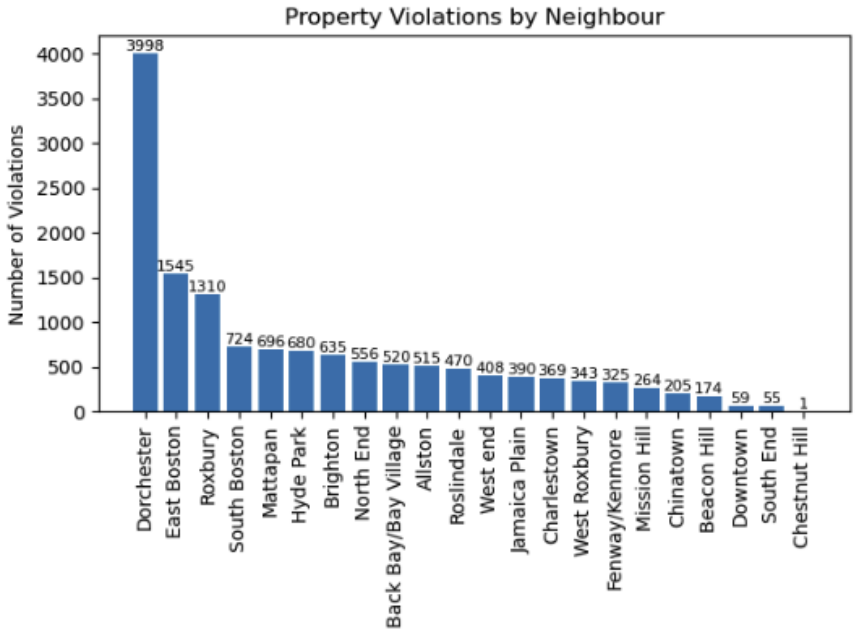
Heat map of property violations(left) and heat map of 311 calls per area(right)

Table 8



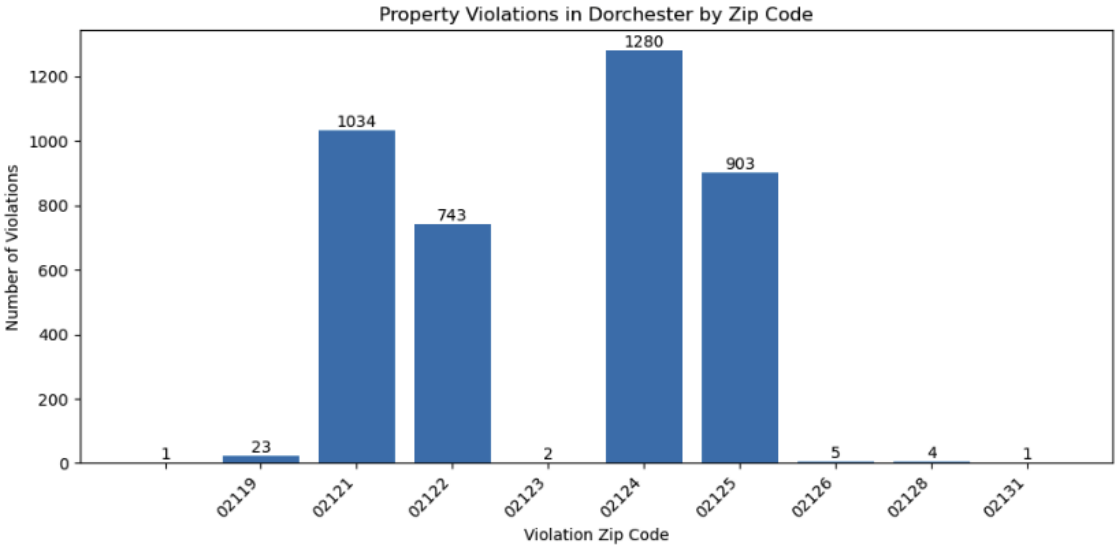
Number of violations in a pie chart

Table 9



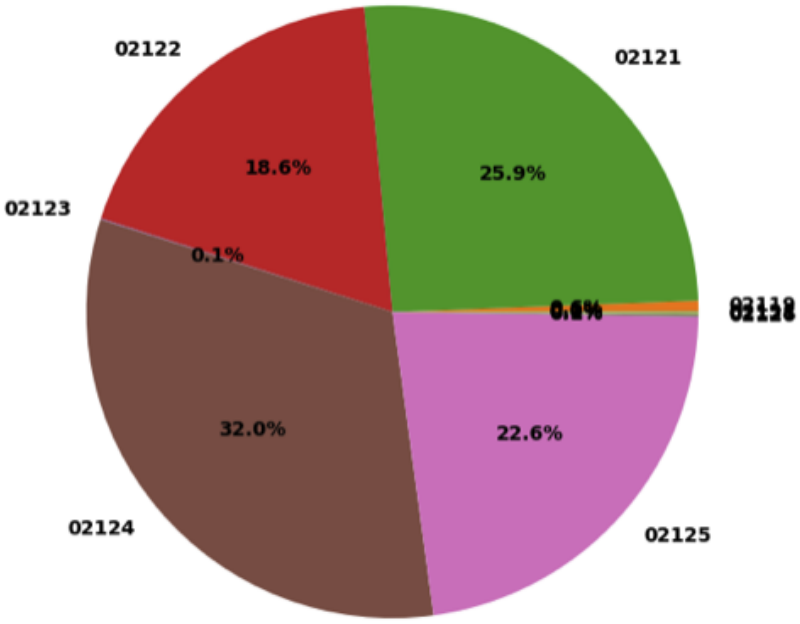
Number of violations in a bar chart

Table 10



Property violations per zip code

Table 11



Most common zipcodes with their property violations.