

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

Client Partner: Sandra Saavedra and Councilor Mejia

Project Manager: Sophia Marian Sena

Technical Engineer: Aidan Gomez

Spark! Staff Lead: Michelle Voong

Team 3 Members:

Derek Dumouchel - Team Lead ddumouch@bu.edu

Zihao Shen zhaojun@bu.edu

Tian Tan tiant@bu.edu

Lu Yao ly21@bu.edu

1. Background & Motivation

Because of the devastating Coronavirus directly resulting in economic uncertainty, the city of Boston made \$12,000,000 investments in various programs, as part of a reallocation of existing city resources to equity and inclusion efforts. This project is organized by Boston City Council and is based on that citywide resilience strategy. The goal is to show where money is going and whether those benefits are being distributed equitably, which can help the Boston City Council assess whether the plan is what they predicted and make adjustments accordingly if the plan deviates.

2. Data Collection

In this project, our client and PM have provided us with several datasets. Since there are too many datasets we can use, we chose 5 datasets eventually. According to the goal of this project, Councilor Mejia wants to understand the geographic area of Boston based on census block groups, so we chose Census Dataset. Then we needed to find relationships between business and population distribution, so we chose Business Funding Dataset, Boston Zoning Subdistricts Dataset, Alcohol License Dataset and Food License Dataset. Those datasets contain sufficient information for us to get results.

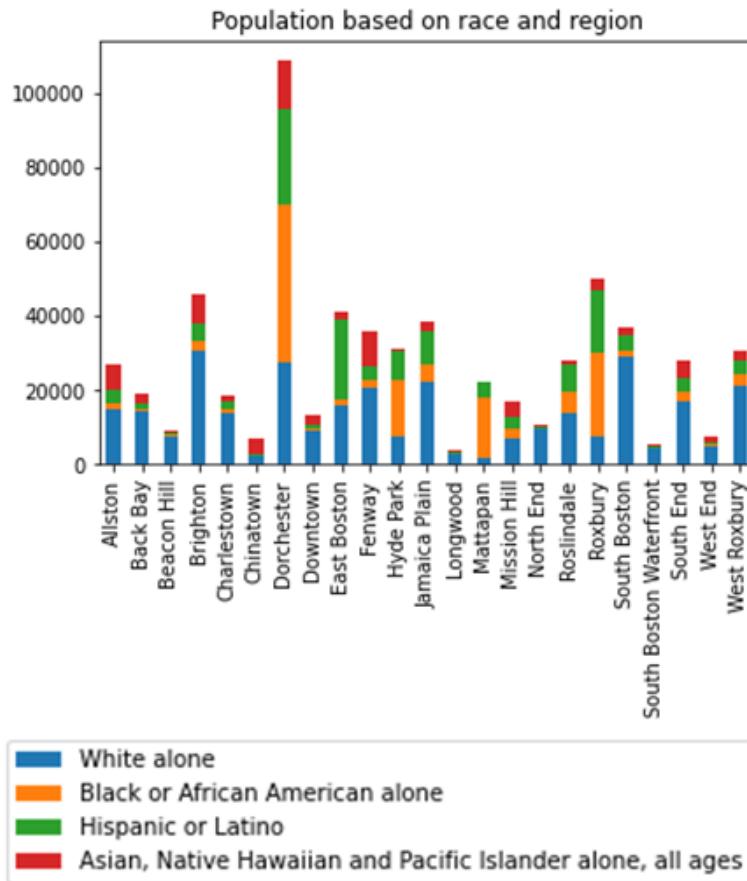
3. Data Visualization and Exploration

After collecting the data set, we performed preliminary data analysis. We explored a total of 4 different datasets including census, business funding, alcohol license, food license. These data sets have many samples, and each sample has many features, but there are some missing values in these features. After we have processed the missing values, we have performed the following visualization and analysis of the data.

Census Dataset.

For the demographic profiles of Boston, we draw a stacked bar chart to show it.

Fall 2022 Councilor Mejia x City Services Project
Team 3 - Final Report



In terms of population, Dorchester is the most populated neighborhood with more than 100,000 people and the ethnic distribution is relatively even. Overall, Asians, Native Hawaiian and Pacific Islanders are a minority group, and only in Chinatown do they account for more than other races. In addition, some of them live more in Fenway. Hispanics or Latinos are mainly distributed in Dorchester, East Boston and Roxbury. Dorchester, Hyde Park, Mattapan and Roxbury concentrated more Blacks or African Americans. The distribution of the White ethnic group is relatively even, and in Back Bay, Beacon Hill, Brighton, South Boston and West Roxbury they account for the vast majority.

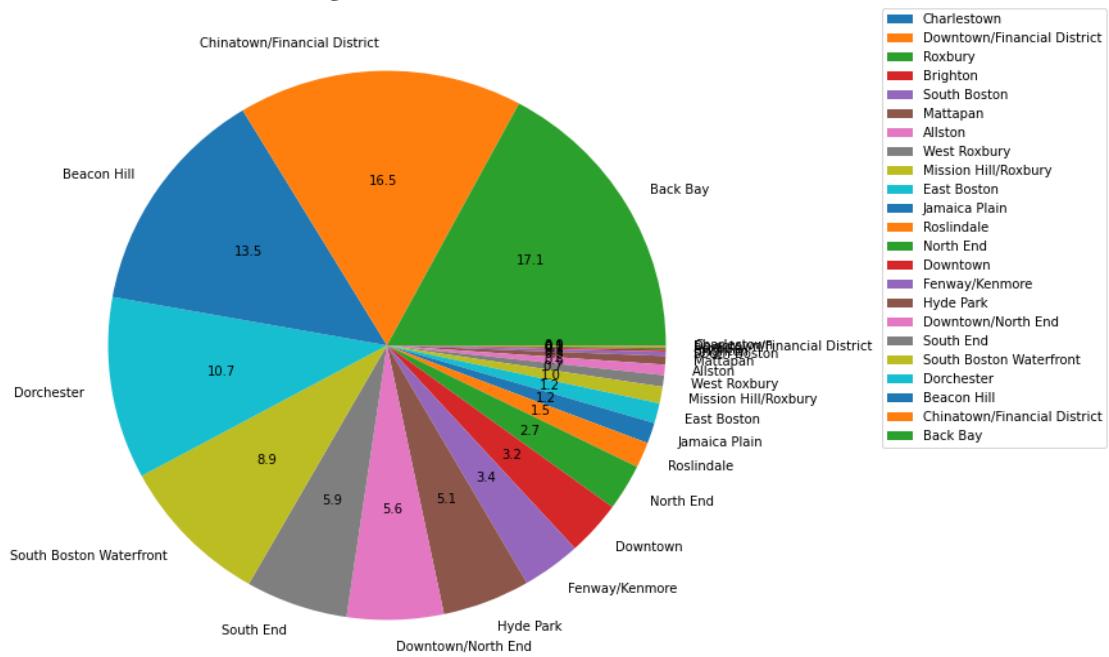
Business Funding Dataset.

Next we plotted a bar chart and a pie chart for the distribution of business funds and sorted them. Here we are displaying a pie chart which more succinctly shows the density of business funds.

From the figure, we can easily conclude that most of the funds flow to the 5 neighborhoods of Back Bay, Chinatown/Financial District, Beacon Hill, Dorchester and South Boston Waterfront. They take around 67% of business funds.

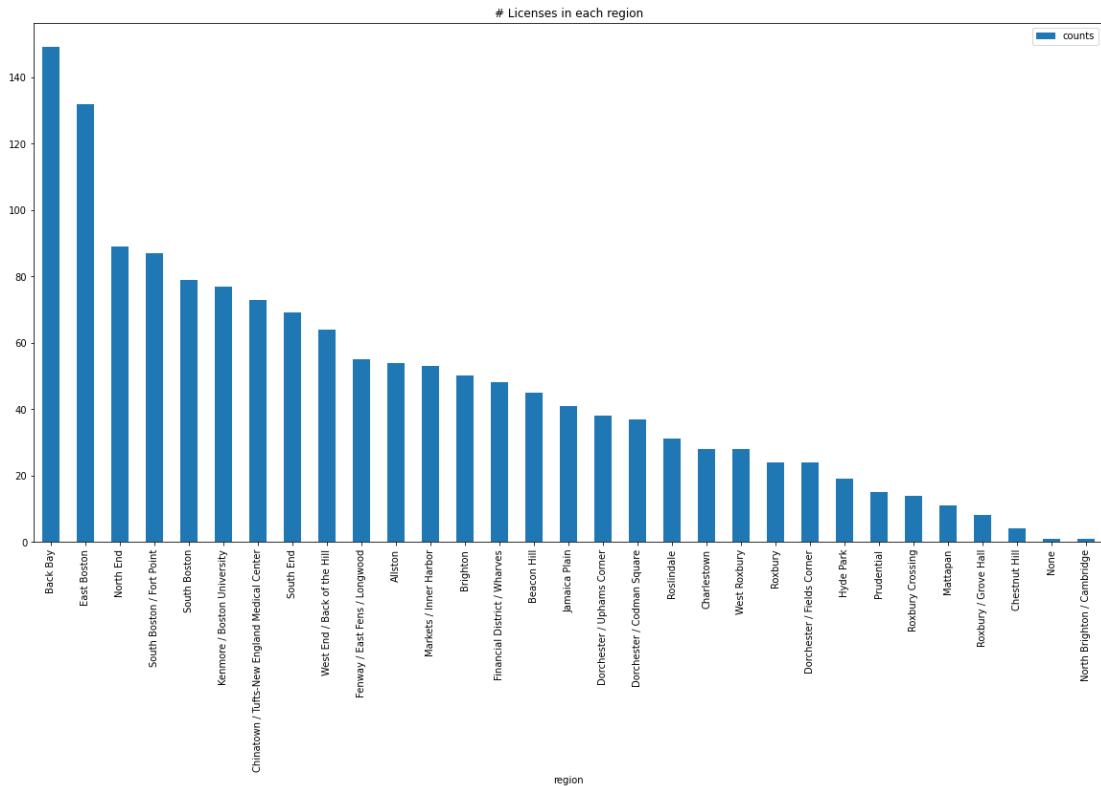
Fall 2022 Councilor Mejia x City Services Project
Team 3 - Final Report

Density of business



Alcohol License Dataset.

Next, we analyzed the alcohol license dataset. We used a bar chart to count the number of licenses per neighborhood. At the same time, we created a confusion matrix of different neighborhoods and different licenses.



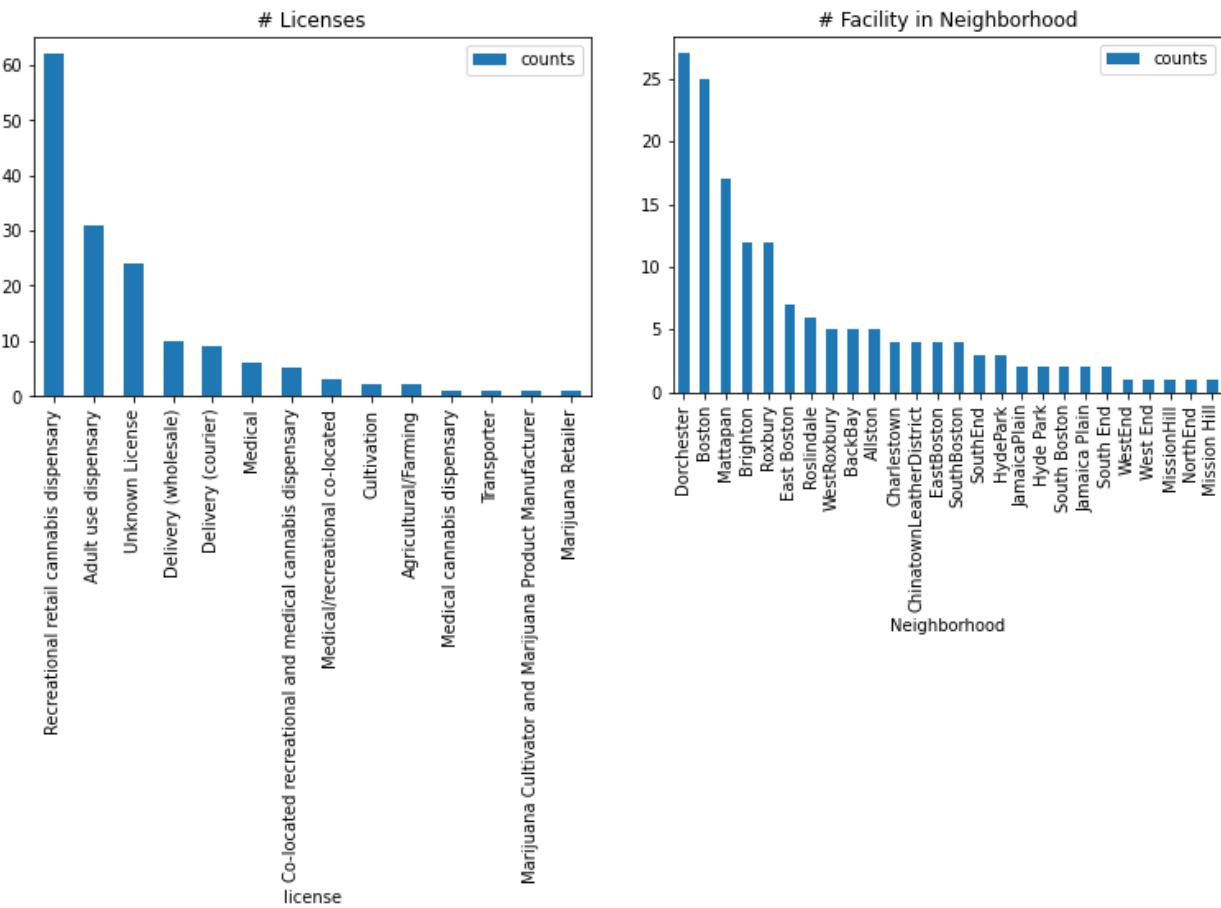
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

From the figure, we can see that Back Bay and East Boston have the largest number of alcohol licenses, more than 120. It is followed by North End, South Boston / Fort Point and South Boston, and their number of licenses are around 80.

Cannabis License Dataset.

For the cannabis license dataset, we analyzed the number of different types of licenses in total and the number of facilities in each neighborhood. The most commonly used license is Recreational retail cannabis dispensary, and then Adult use dispensary. About 15% of the records didn't show the type of the license. There are not many facilities. Dorchester and Boston downtown have the most facilities, then Mattapan, then Brighton and Roxbury.

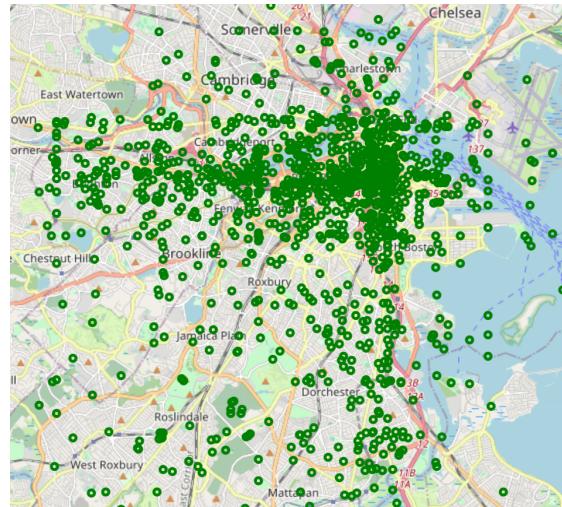
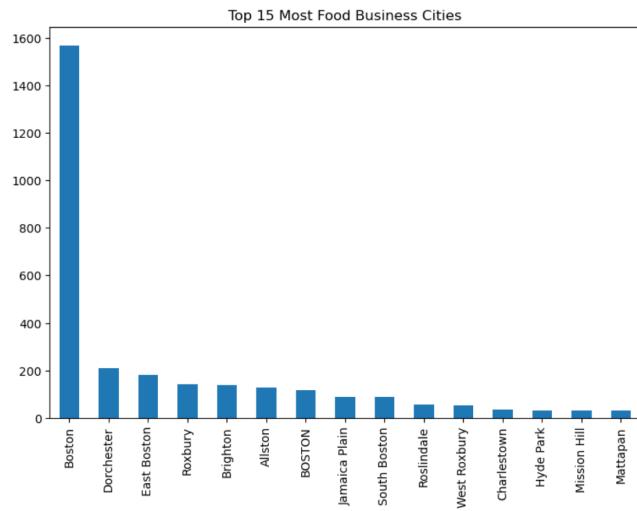


Food License Dataset.

For the food dataset we plotted a bar chart listing the top 15 neighborhoods with the most food licenses. Additionally, we have aggregated regions by number of licenses and created a pie chart. Finally we marked the food business on the map.

Fall 2022 Councilor Mejia x City Services Project

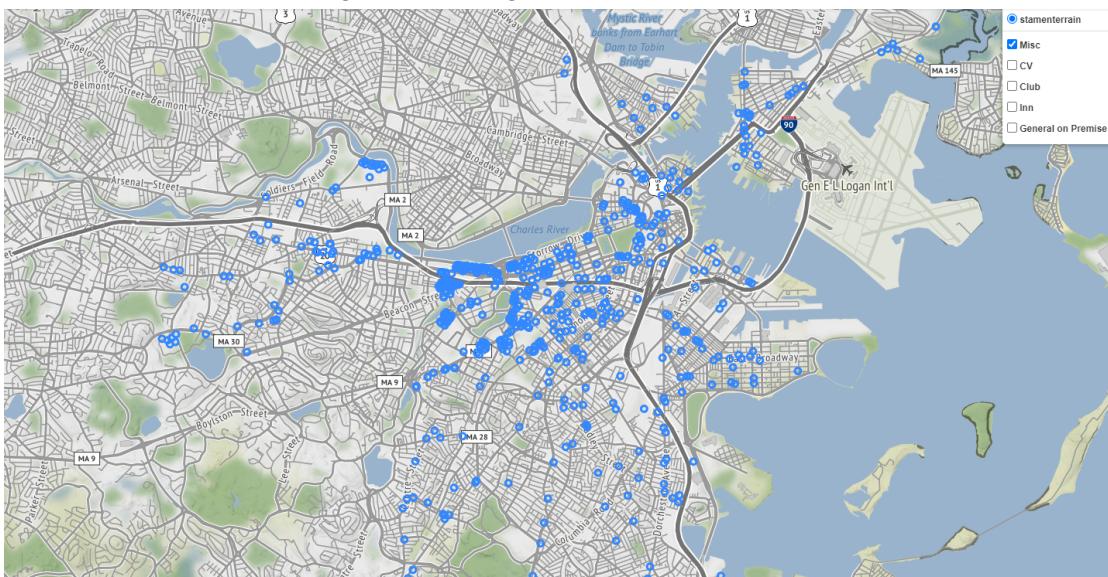
Team 3 - Final Report



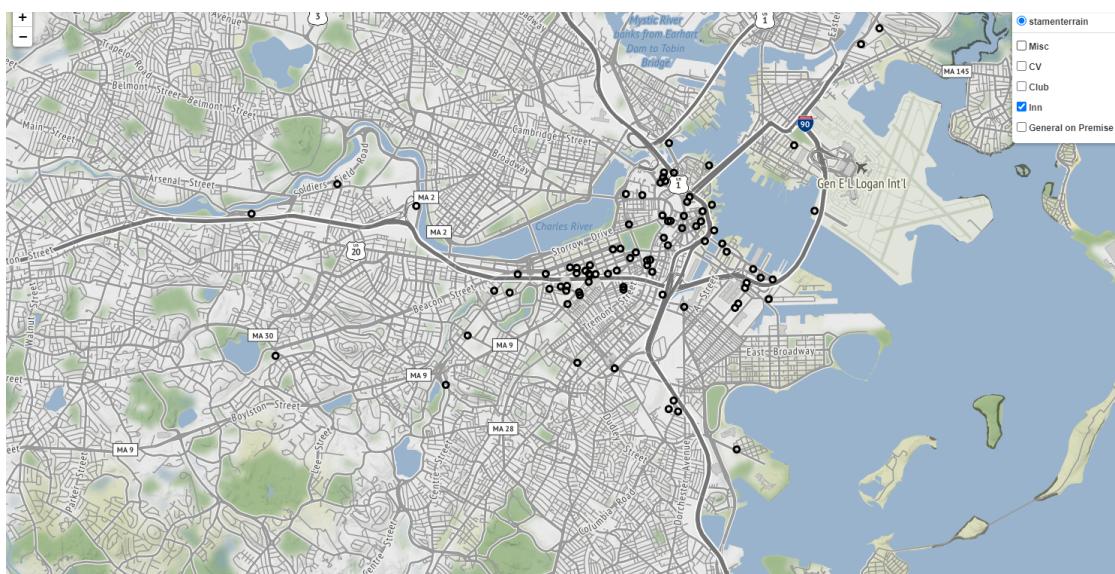
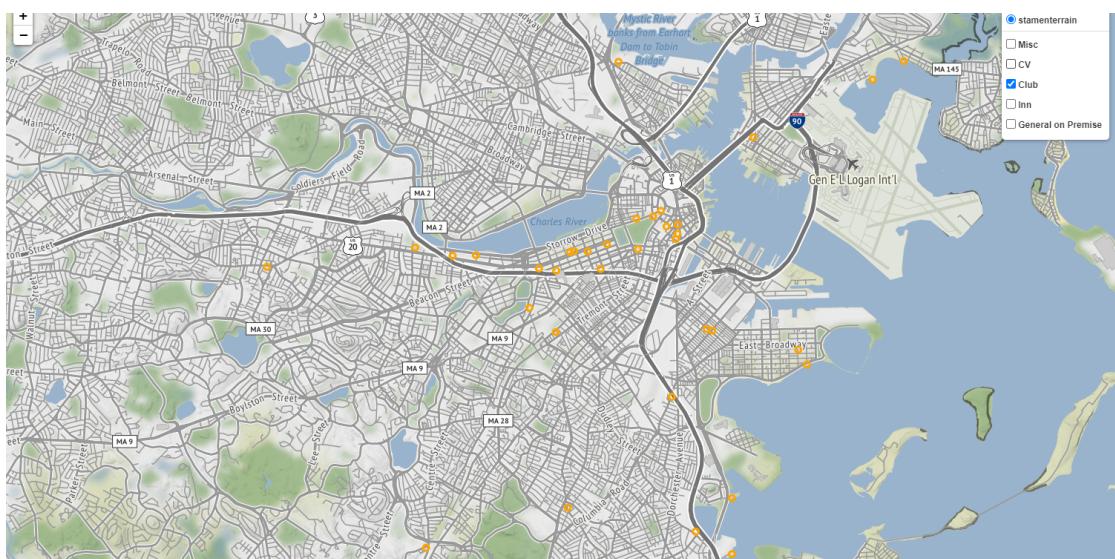
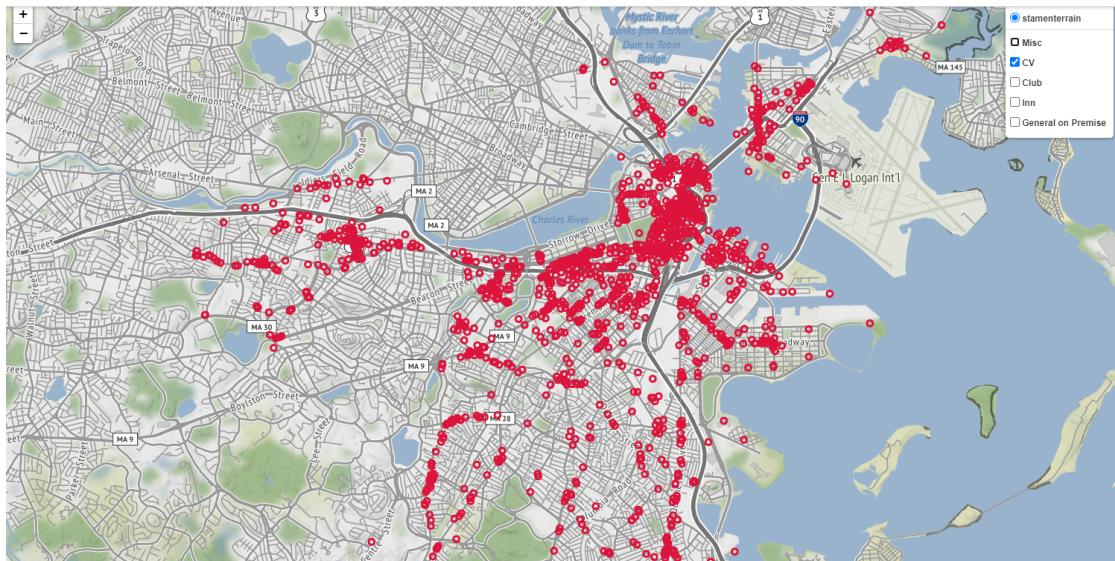
Bar chart shows the number of food businesses in each city. It is very easy to find that the number of food businesses in Boston far exceeds that of other regions. In other regions, the food business is more evenly distributed. From the pie chart which doesn't display here because of space, cities with less than 30 food businesses account for 43.8% of all cities. On the map, it shows that their distribution is scattered around the Boston city.

Licensing Board Licenses Dataset.

This dataset lists a variety of currently active license under MGL Chapter 138 and Chapter140 including Billiards, Bowling, Clubs/Veterans' Groups, Common Victualler (Food Service), Alcohol Beverage, Innholder, Dormitories/Lodging houses, and Retail Package Stores. Those licenses can be divided into five categories: Misc, Common Victualler, club, Inn and General on Premise. We generated a Boston basemap and plotted locations of those licenses on the map with five different markers according to five categories to see the distribution.

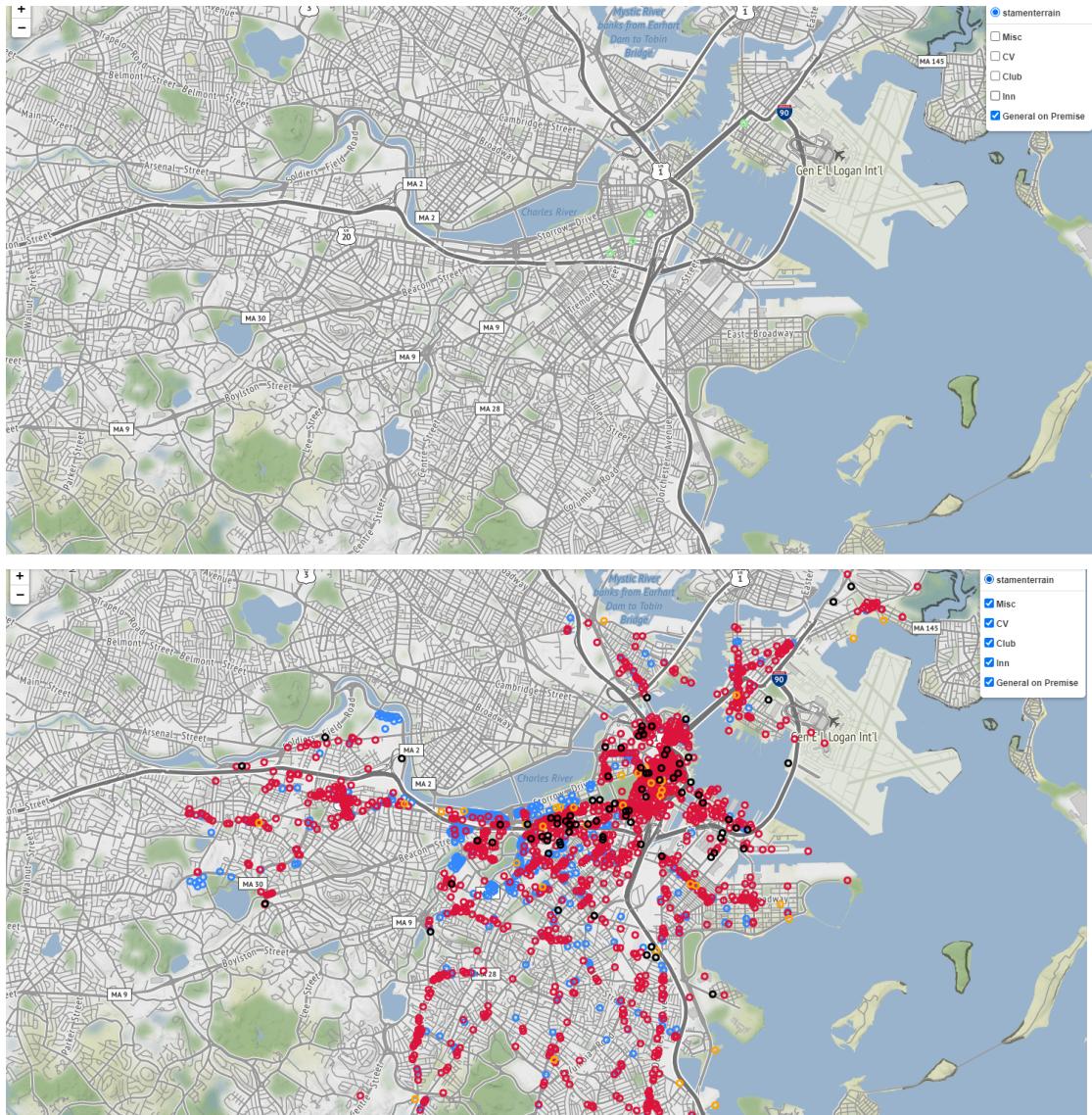


Fall 2022 Councilor Mejia x City Services Project
Team 3 - Final Report



Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report



Rental Assistance Dataset.

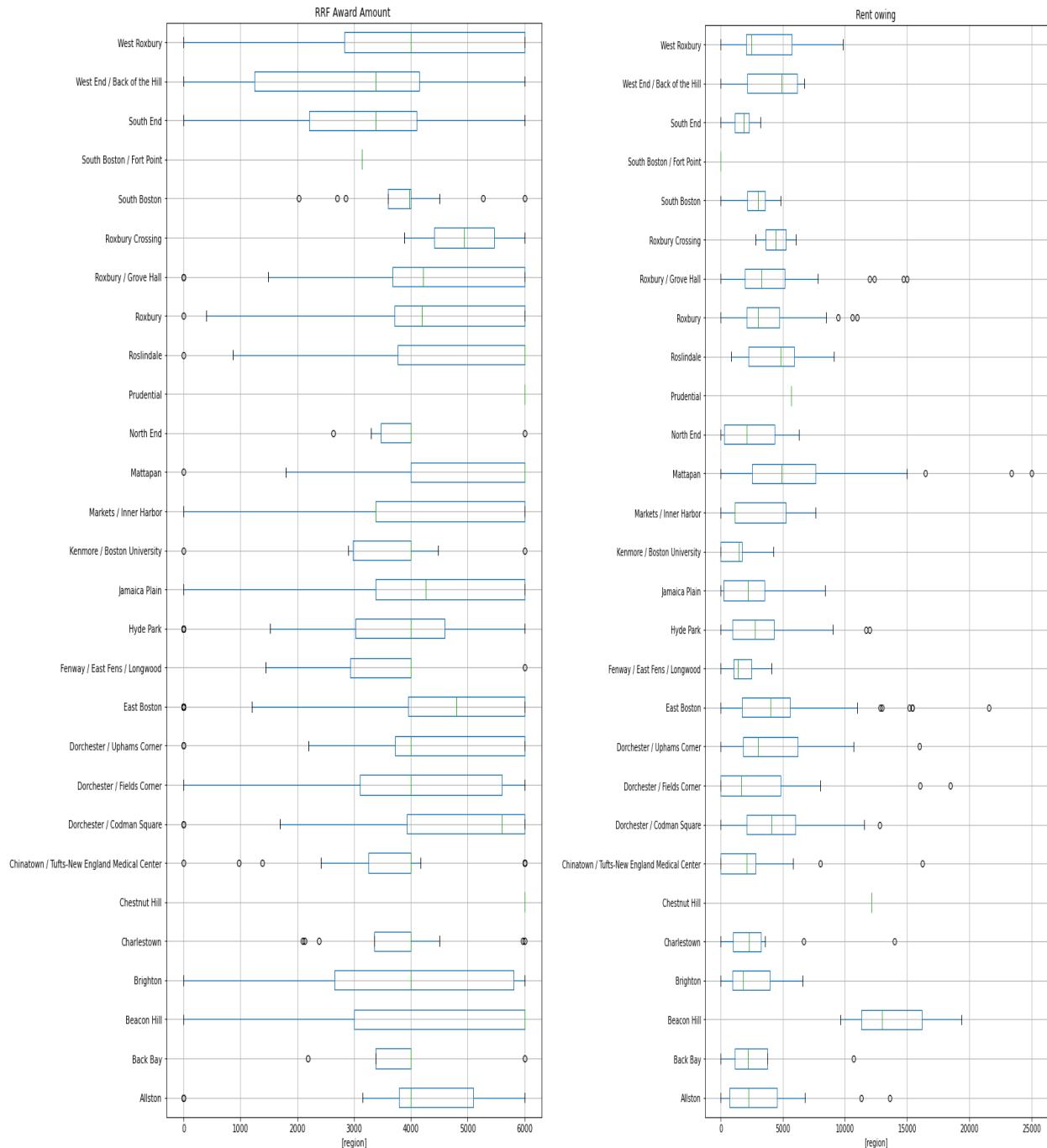
For the rental assistance dataset, we focused on the RRF Award Amount and Rent Owing. We plotted the box plot and bar chart for these two columns based on the neighborhood (zip code).

Firstly, we analyzed the whole dataset.

From the box plot about the RRF Award Amount, we can find that at this moment, the most rental assistance people got is \$6000. And over half of people in Roslindale, Mattapan, and Beacon Hill got \$6000.

Fall 2022 Councilor Mejia x City Services Project
Team 3 - Final Report

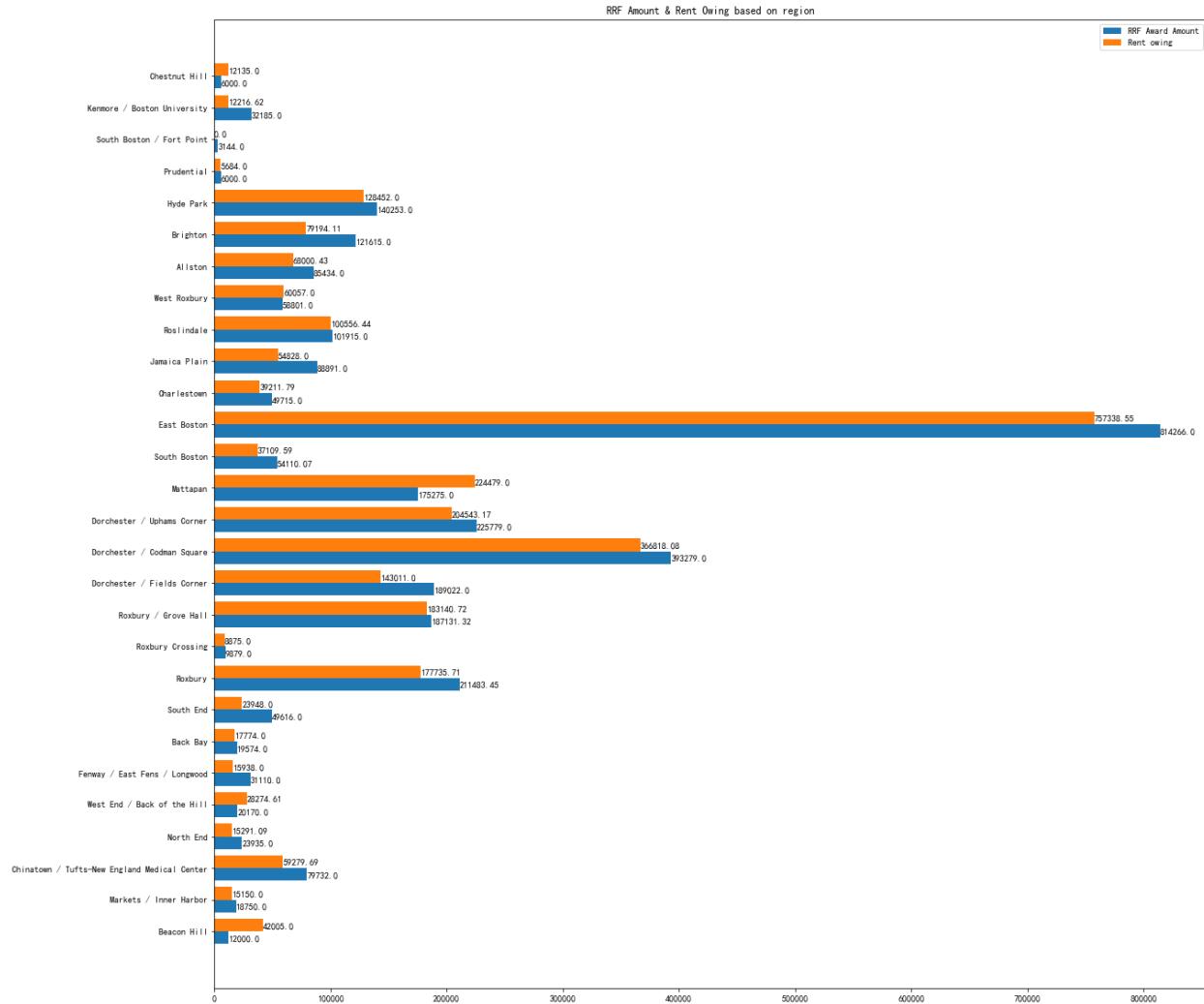
From the box plot about Rental owing, we can summarize that Beacon Hill had the highest amount of rent owing. And South Boston does not have rent owing.



Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

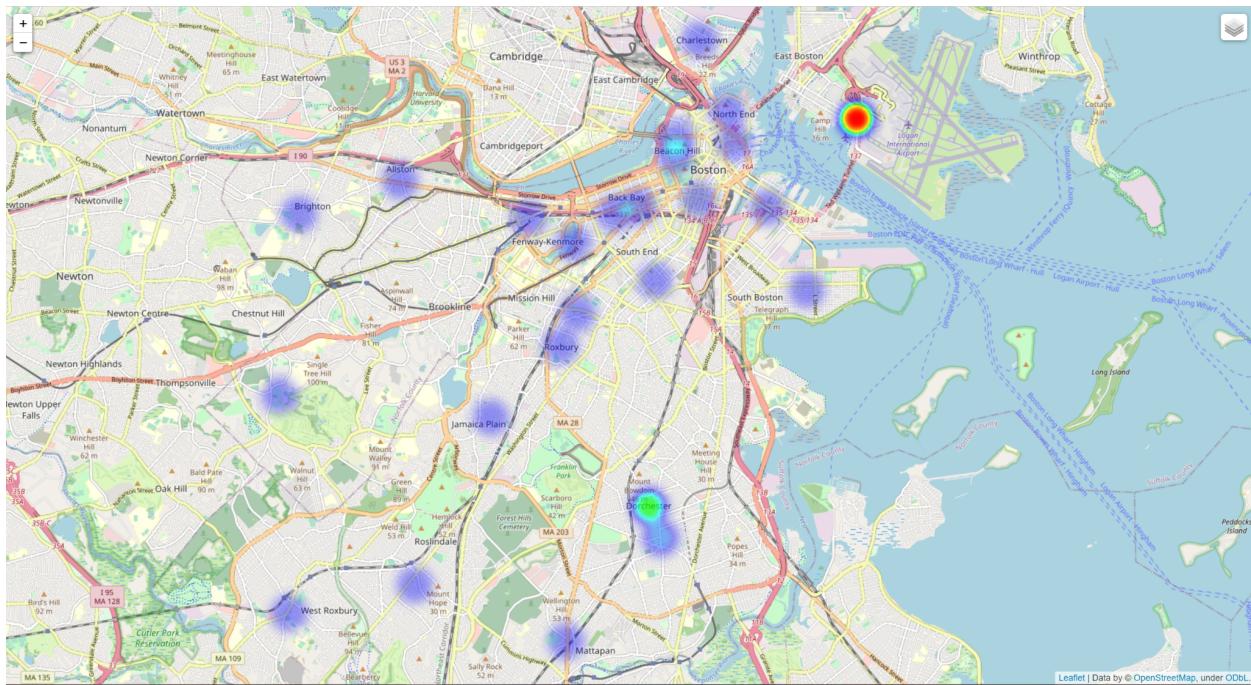
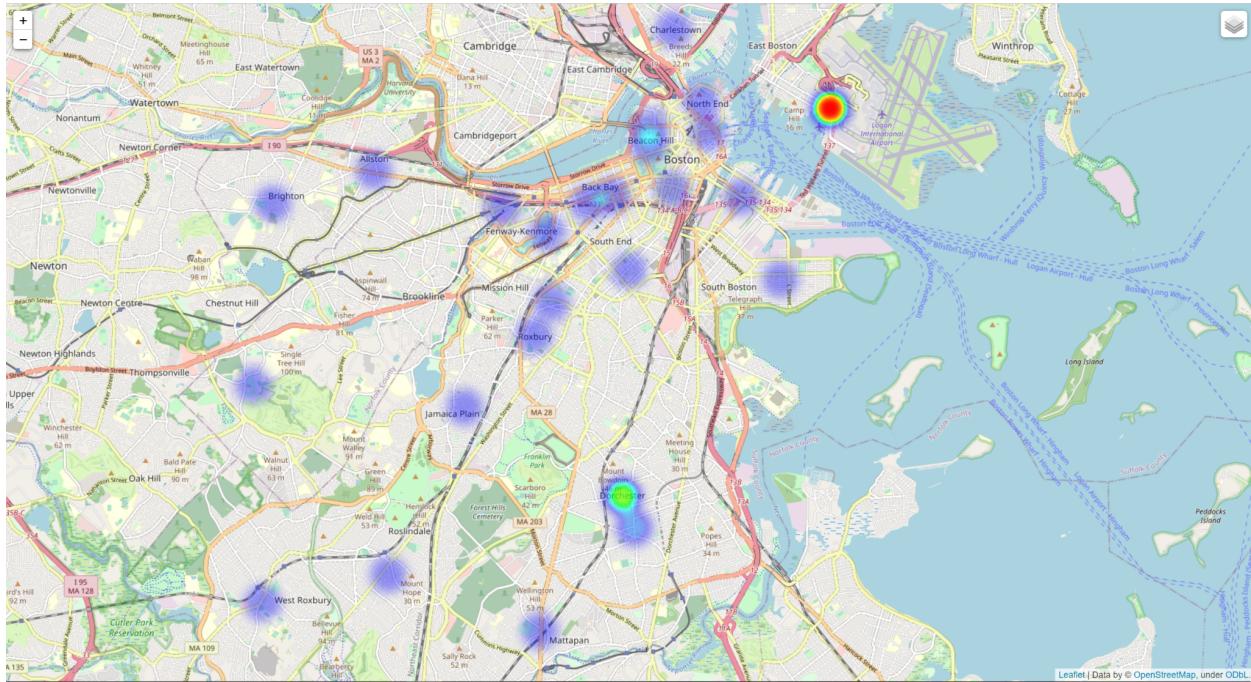
Then we computed the sum of the RRF Award Amount and Rent Owing based on the neighborhood. It is clear that East Boston got the most RRF Award and had the highest Rent Owing. And if we consider the sum of three parts of Dorchester, it got the second highest RRF Award.



Finally, we plot the heatmap based on the sum of the RRF Award Amount and Rent Owing in each neighborhood. It looks similar. And It is clear to show the difference among sum of two types of data in each neighborhood.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report



Capital Investments Dataset.

This dataset is the City's plan to make critical investments in the City's infrastructure. This year's \$3.6B plan includes new infrastructure funds supporting affordable housing, revitalized parks, multi-modal transportation improvements, school improvements across the city.

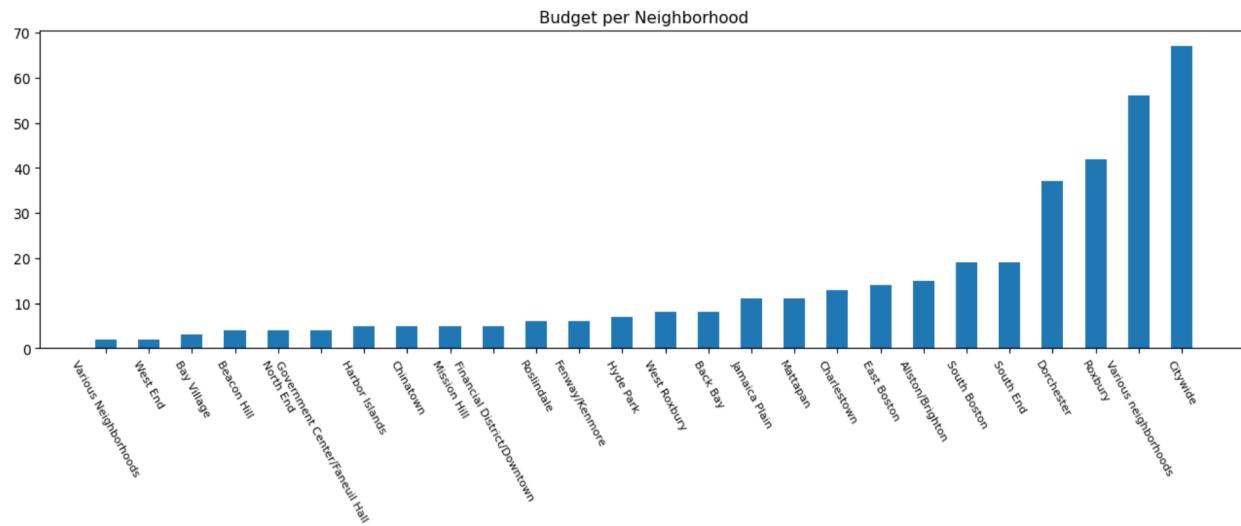
We counted the budget totals for all neighborhoods. It can be found that 'Citywide' has the highest total budget. Meanwhile, the status of all projects is also analyzed. The status of most

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

projects are: In Design, New Project and To Be Scheduled. This can reflect the very good project vitality in Boston, which means that there are many new projects being conceived.

Project_Status	Department			PM_Department			Neighborhood			Total_Project_Budget						
	count	unique	top	freq	count	unique	top	freq	count	unique	top	freq	count	unique	top	freq
Annual Program	46	8	Transportation Department	12	46	9	Transportation Department	12	42	6	Citywide	27	46	42	1,000,000	4
Complete	1	1	Boston Public Schools	1	1	1	Boston Public Schools	1	1	1	Citywide	1	1	1	7,222,010	1
Implementation Underway	28	11	Department of Innovation and Technology	6	28	11	DolT	6	22	4	Citywide	15	28	23	1,000,000	6
In Construction	57	12	Boston Public Schools	14	57	6	Public Facilities Department	27	57	19	Roxbury	9	57	52	1,750,000	2
In Design	86	12	Parks and Recreation Department	24	86	8	Public Facilities Department	28	86	24	Roxbury	12	86	73	1,000,000	6
New Program	1	1	Parks and Recreation Department	1	1	1	Parks and Recreation Department	1	1	1	Citywide	1	1	1	500,000	1
New Project	86	15	Boston Public Schools	22	86	14	Public Facilities Department	39	77	20	Roxbury	12	86	46	500,000	12
Study Underway	14	6	Parks and Recreation Department	5	14	4	Public Facilities Department	8	12	7	Various neighborhoods	3	14	12	200,000	2
To Be Scheduled	88	15	Parks and Recreation Department	21	88	14	Public Facilities Department	33	80	20	Various neighborhoods	14	88	50	1,000,000	8



Main Streets Business (New dataset provided by client in 12/2)

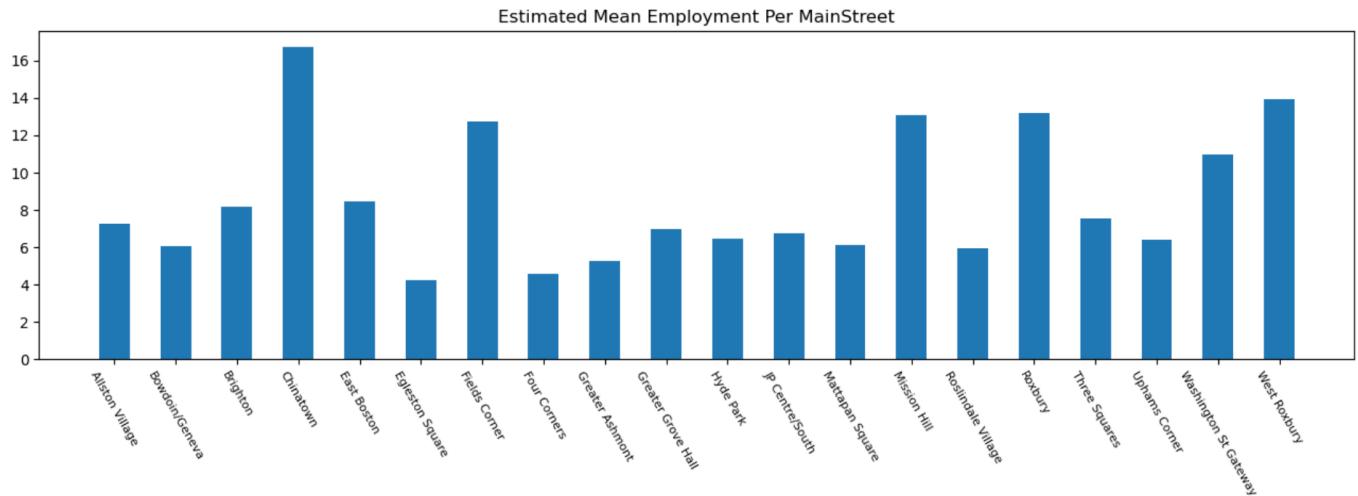
It's a new aggregated mobility dataset of individuals going to/from points of interest across businesses.

First we analyzed the Estimated Mean Employment Per Main Street. We find that Chinatown and West Roxbury have the largest estimated average employment numbers, indicating that these two streets are more densely populated and have more commercial activities. This is in line with our results from other datasets.

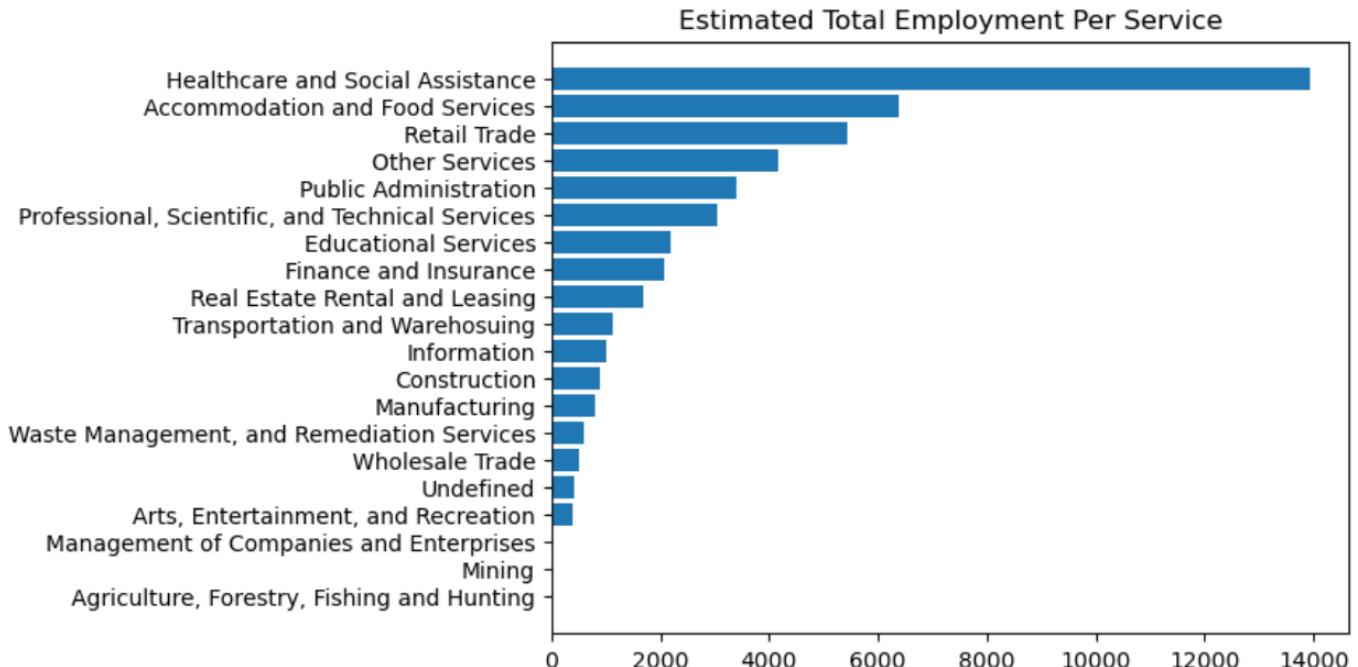
Then we analyzed the Number of Business Services Per Main Street and Estimated Total Employment Per Service. For the analysis of the number of business services per street, we can get the same results as the previous analysis. Chinatown is still the street with the most business services. We plotted a bar chart of it in code which is not shown here.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report



The chart below shows the estimated number of employment in each business service. It is very obvious that the healthcare service requires the most employment, it is twice as much as the accommodation service. This is very logical. In the context of the pandemic, it accurately reflects the current living needs of citizens. Citizens are under increasing pressure on medical care and accommodation.



4. Results Obtained / Questions Answered

• Results

This time, we analyzed all available datasets provided by the base project. It contains 5 categories of data, a total of 7 datasets.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

The number of licenses in each area can show whether this area needs more funds. Then we can compare figures we plotted to find the potential connections among them.

From license datasets, we can conclude that Back Bay, Financial District, Beacon Hill, and South Boston Waterfront have more licenses than other areas.

At the same time, from the business fund dataset, we can see that the following places: Back Bay, Chinatown/Financial District, Beacon Hill, Dorchester and the South Boston Waterfront accounted for 67% of the funds. Thus the funds of the council were basically delivered to them.

Thus, from this perspective, it is reasonable to say the city council's funds are being equitably distributed. Because communities with a higher number of business licenses indicate more economic activity, it is reasonable to allocate more council funds to these areas.

In addition, we can also draw a conclusion that in the context of the pandemic, people's demand for accommodation and healthcare has greatly increased and through the capital investment dataset, it can be found that most of the funds have flowed to New projects or In Design projects which aim to support affordable housing.

Therefore, it is reasonable to say that municipal funds are distributed fairly and that funds go where they are expected to go.

• Answer questions

1. Where did business assistance go during the pandemic? What were the demographic profiles of the communities where the businesses were located?

Answer:

- Most of the funds flow to the 5 neighborhoods of Back Bay, Chinatown/Financial District, Beacon Hill, Dorchester and South Boston Waterfront. They take around 67% of business funds.
- Among them, Dorchester is the most populated neighborhood with more than 100,000 people and the ethnic distribution is relatively even. Asians account for more than other races in Chinatown. Besides, the White ethnic group accounts for the vast majority in the remaining 3 areas.
- Visualization and graphs are shown above.

2. Where did the city's rental assistance go during the pandemic (Average is 4532, Chestnut Hill and Prudential get the most rental assistance)? What were the demographic profiles of the recipients of these funds?

Answer:

- East Boston (\$814266.0), Dorchester (\$808080.0) got the most rental assistance. They take around 50.5552% of rental assistance.
- Based on the race analysis, in the two regions, White, Black or African American, Hispanic or Latino take the similar percentage, and Asian or Native Hawaiian or Pacific Islander take the least percentage.
- Visualization and graphs are shown in part 8.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

3. Where are the city's economic development licenses? Which communities are benefitting? Which communities are being left out?

Answer:

- The city's economic development licenses are mainly in Boston City.
- Back Bay and East Boston benefit most, while Charlestown, Hyde Park, Mission Hill and Mattapan are being left out.
- Visualization and graphs are shown above.

5. Limitations of Results

Our results are extrapolated by combining and comparing results from several different datasets. In addition, most of our analysis is based on region, and it is not possible to count the percentage of funds assistance received by each race. Also there are many missing values in the datasets.

6. Challenges Faced

- Datasets are messy and ambiguous. For example, for the census analysis part, we are provided with 3-4 different population data sets, each data set is similar but has some unique characteristics, we need to analyze each data set and filter out some datasets. Some of these datasets are not up to date. It is difficult for us to determine which dataset is the most suitable.
- Data's categories are various, it is hard to combine them together. The goal of our project is to explore whether city council assistance funds are being allocated properly. But we are provided with census dataset, enterprise business funds dataset, and business license dataset. There is no real dataset of funding allocations. In this way, we have no real data to test and verify our conjecture or result.

7. Suggestions for the Future of the Project

We need to further analyze the housing assistance dataset, explore its connection with other datasets more deeply.

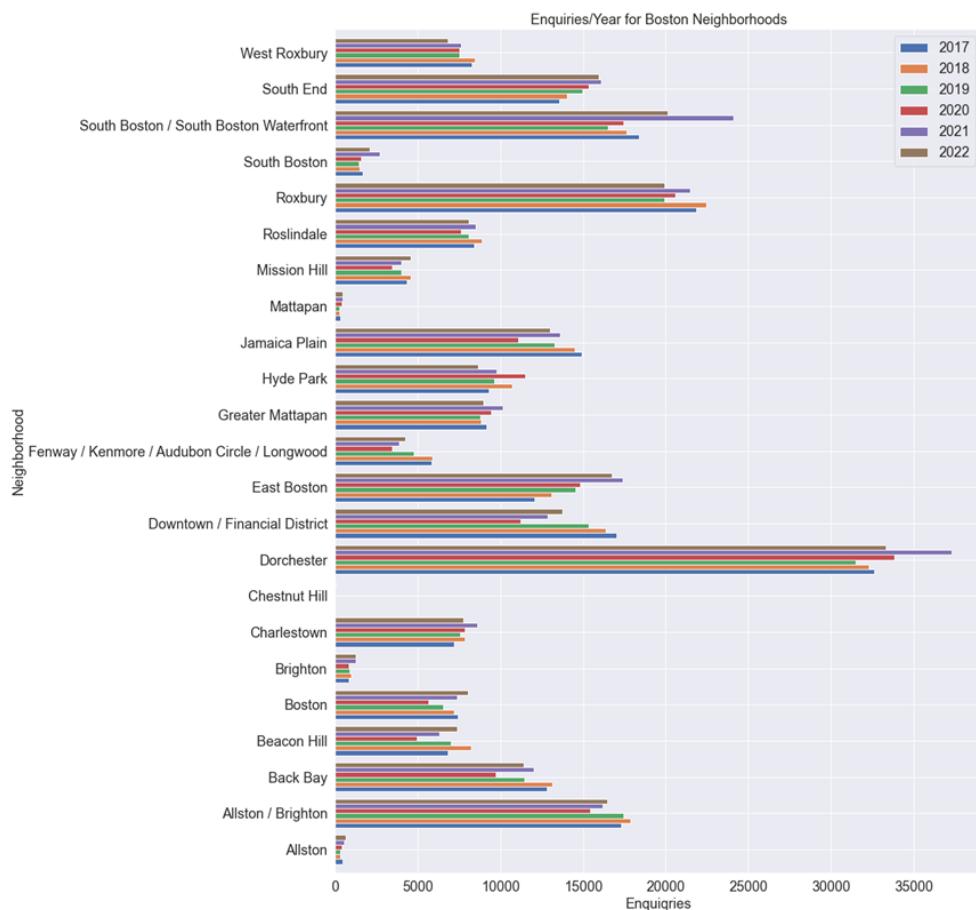
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

8. Extension Project

Initially, the team's intentions were to further explore the rental assistance and housing assistance datasets as part of the extension project. An interesting topic that the team was aiming to study was the effect of the Covid-19 pandemic on applicants for rental and housing assistance. The intent was to review first-time home ownership data for pre- and post-pandemic applicants. Additionally, it was desired to further the base project's evaluation on whether relief funds were proportionally and adequately distributed across neighborhoods. The 17F Request RRF dataset was intended to aid with the extension project, but contained roughly 70% ineligible requests. Due to the limited availability of rental and housing assistance data the decision was made to access the Boston 311 datasets for the extension project.

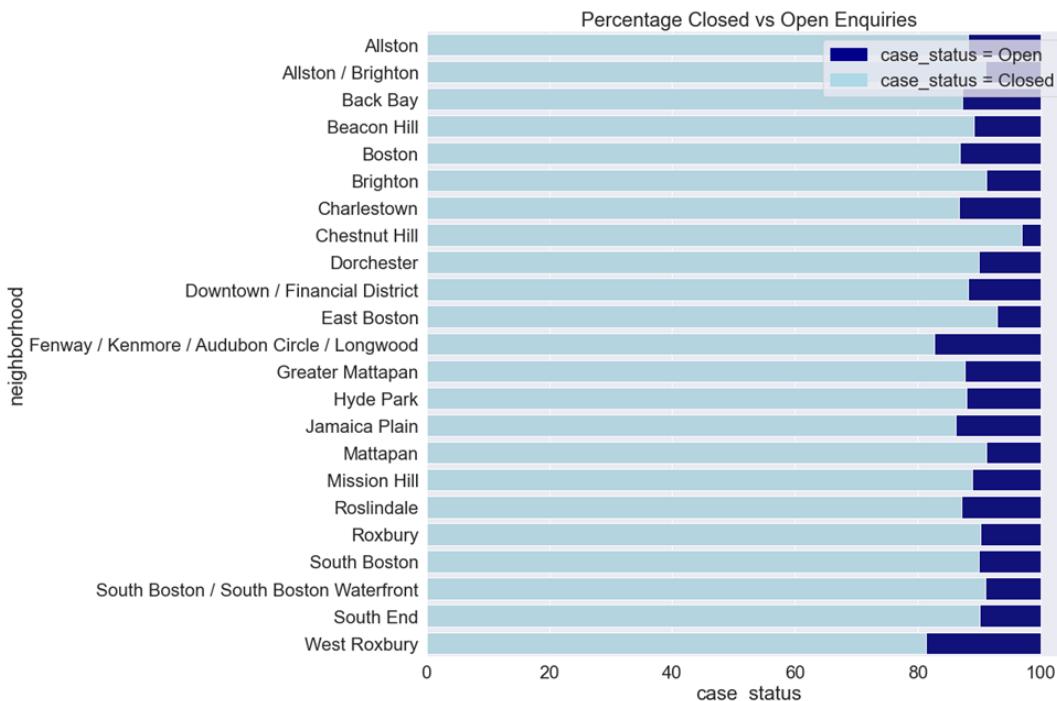
The Boston 311 call service connects users to service center representatives to log caller requests for city services. Each service request is logged with a specific subject, reason and type along with the service open date, location, and current closure status. Service request datasets are available from 2011 to present day. A broad look at enquiries across Boston neighborhoods was performed from 2017 to present. The total number of enquiries for each Boston neighborhood is shown in the figure below and shows the total number of enquiries is generally consistent throughout the years.



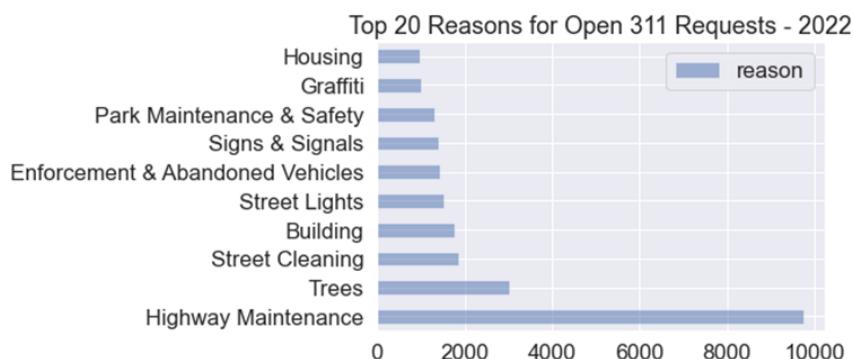
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

Refined assessments for breakdowns of enquiries were performed using data from 2022 to keep visualizations reasonable and readable. The scope of the 311 extension project is to evaluate the types of 311 enquiries by neighborhood and their typical closure times. By examination of these enquiries, the influence of neighborhood demographic on service times can be evaluated. First, neighborhoods were compared to see the number of closed versus open enquiries. As shown in the figure below, the proportion of closed enquiries to open enquiries is rather consistent across neighborhoods. This is likely because the year-end is nearing.



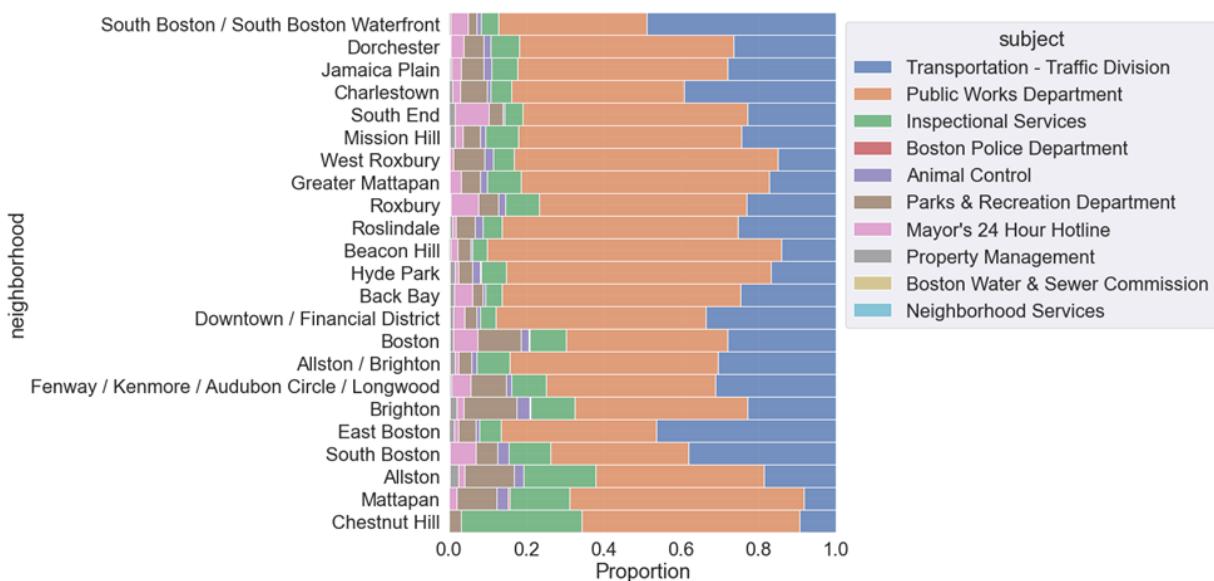
The plot below shows the top reasons for enquiries in 2022 that have not been closed. Most open enquiries are related to highway maintenance, which is expected as infrastructure repairs typically take longer than say, street sweeping.



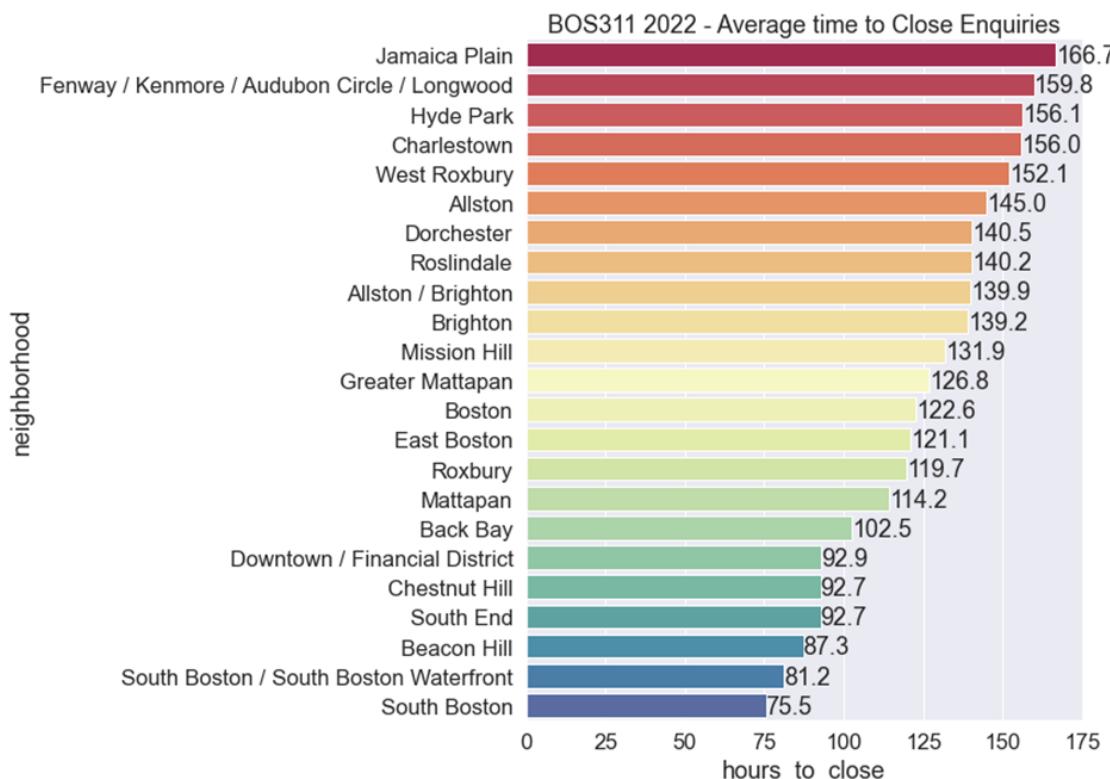
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

Of the closed inquiries, the most common for neighborhoods were assigned to the Public Works Department or the Transportation – Traffic Division. A breakdown of the proportion of enquiries is shown below.



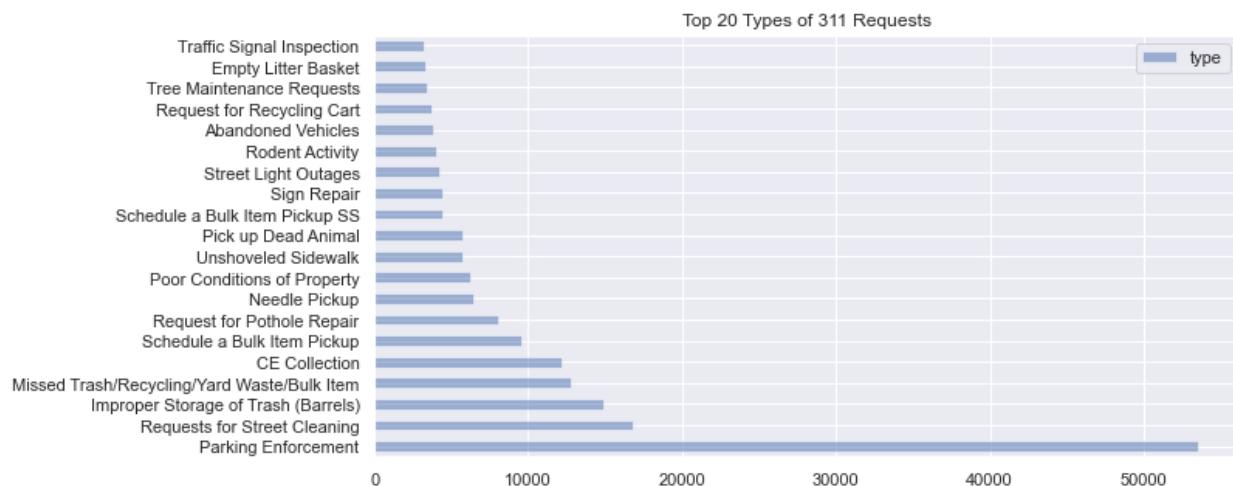
The figure below shows the average closure times by neighborhood, with Jamaica Plain, Fenway area, and Hyde Park, taking the longest and South Boston, South End, and Beacon Hill taking the shortest time to close.



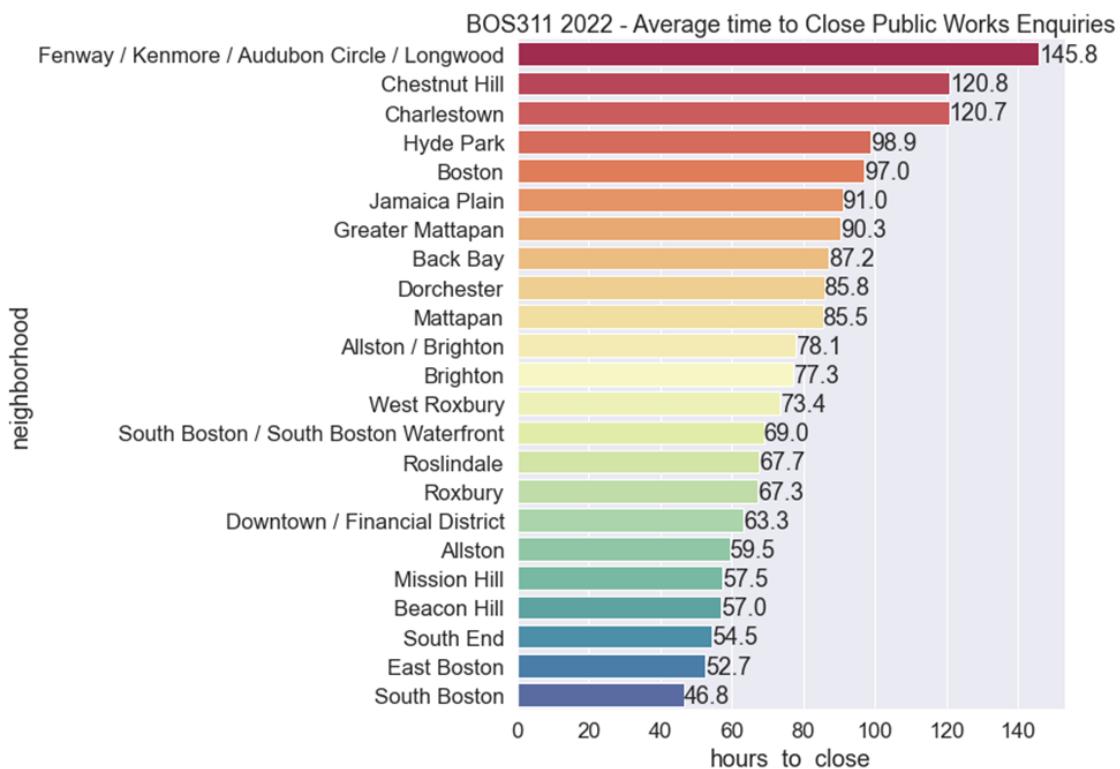
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

The top types of requests are plotted below and are overwhelmingly related to parking enforcement followed by street cleaning and improper storage of trash barrels.



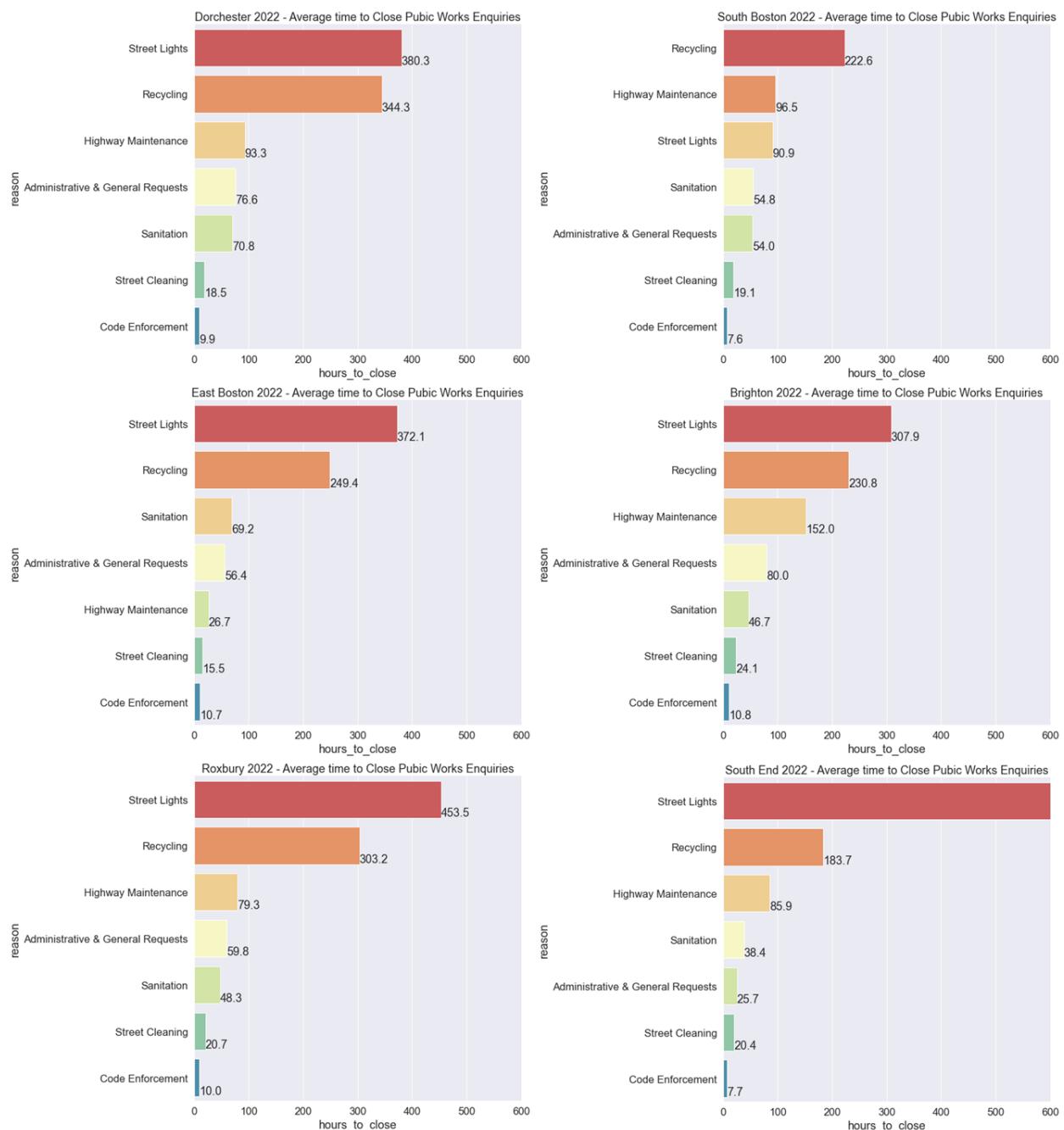
It is possible longer lead issues not related to public works may affect lead times. Examination of only Public Works Enquiries shows a similar trend in neighborhood to response times, and is shown in the figure below.



Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

From the base project, it was shown that the neighborhoods with the largest percentage of Black and Hispanic and Latino populations include Dorchester, East Boston, and Roxbury. Neighborhoods with a larger percentage of white population include South Boston, Brighton, and the South End. The figure below plots the breakdown of public works enquiries by the above neighborhoods and their average time to close. It is observed that there is no direct link between neighborhood race diversity and response time of enquiries.



Breakdowns of the same neighborhoods by type of enquiry are plotted in the figure below.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Final Report

