Fall 2022 Councilor Mejia x City Services Project

Team 3 - Draft 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	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

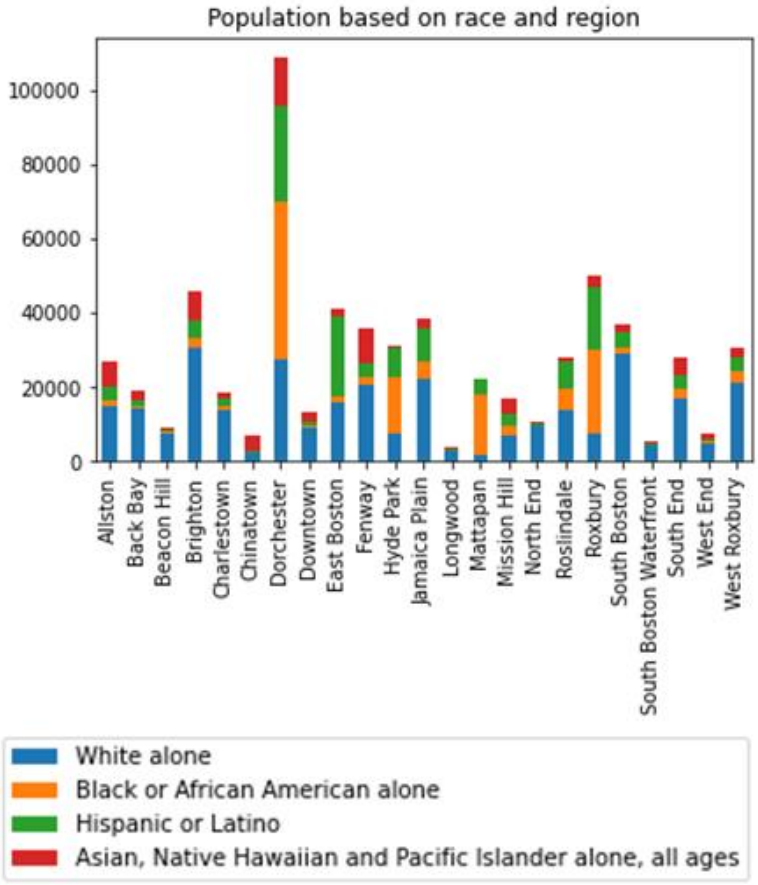
After collecting the data set, we performed preliminary data analysis. We explored a total of 5 different datasets including census, business funding, zoning subdistricts, alcohol license, food license. These data sets have many samples, and each sample has many features, but there are some missing values 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 - Draft 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.

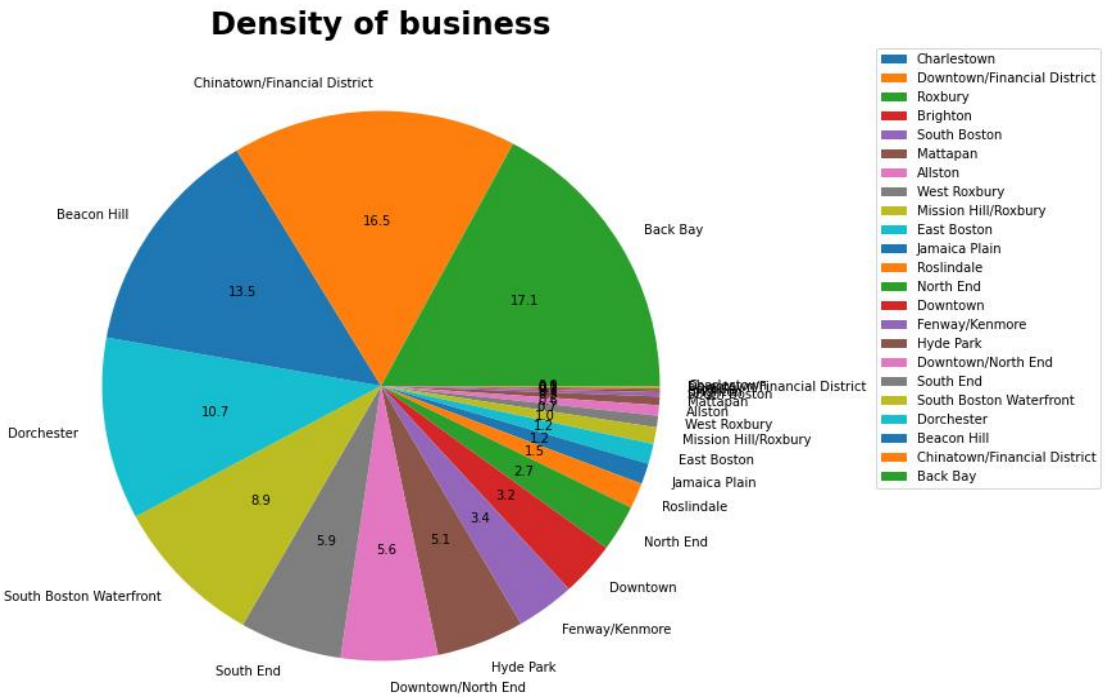
Business Funding Dataset and Boston Zoning Subdistricts Dataset.

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.

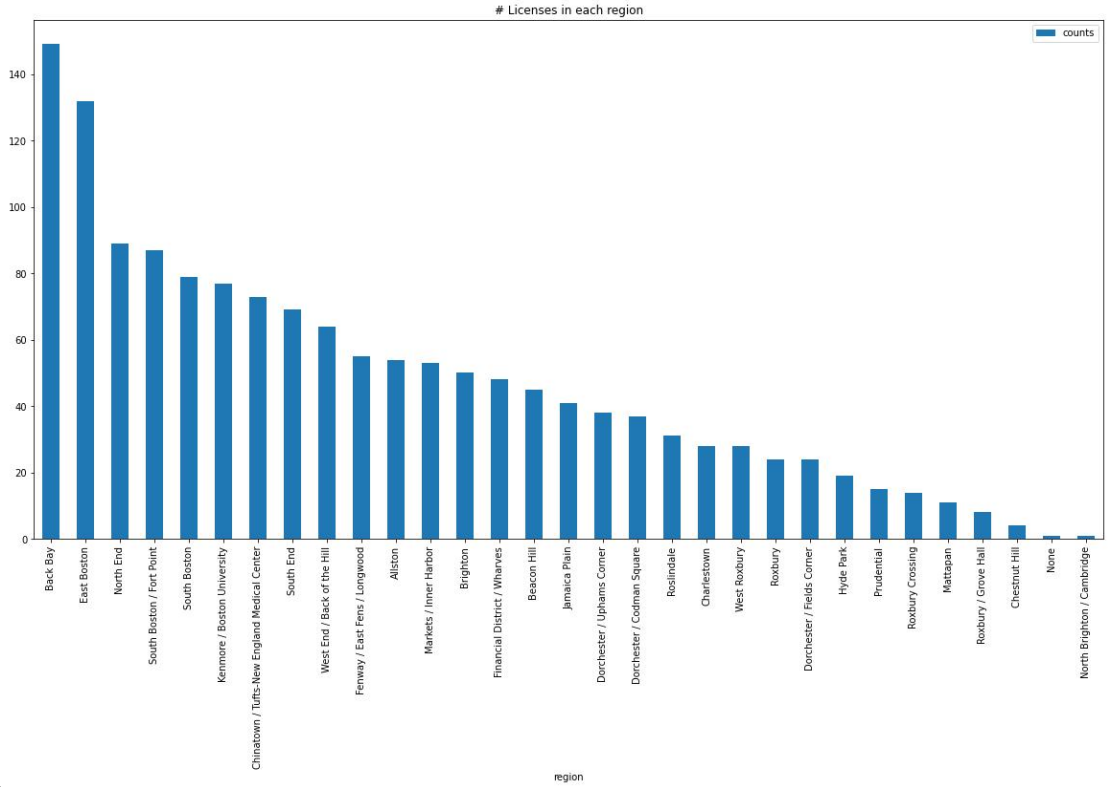
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Draft Report



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

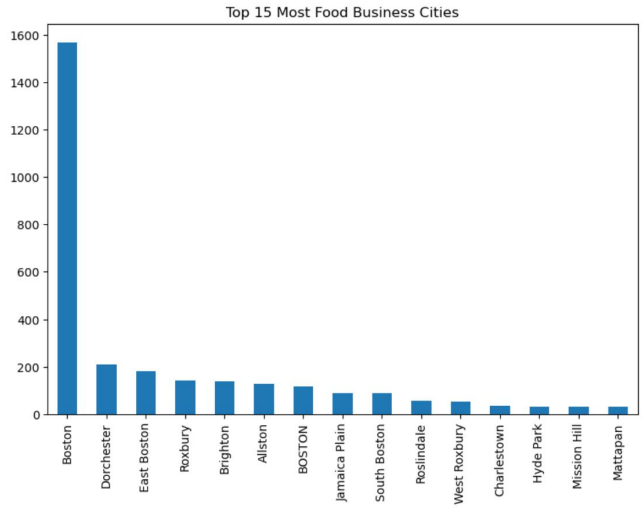


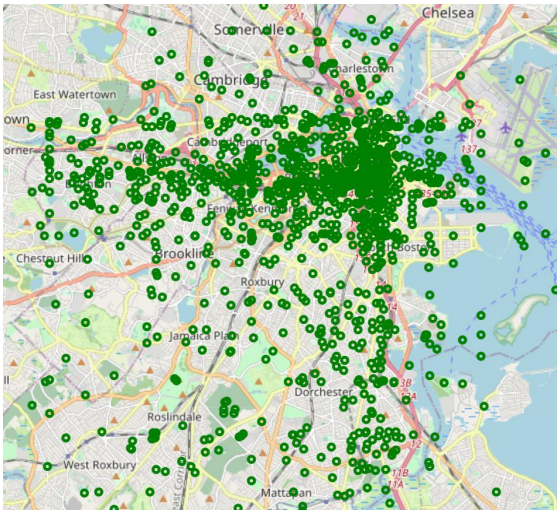
licenses.

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.

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.





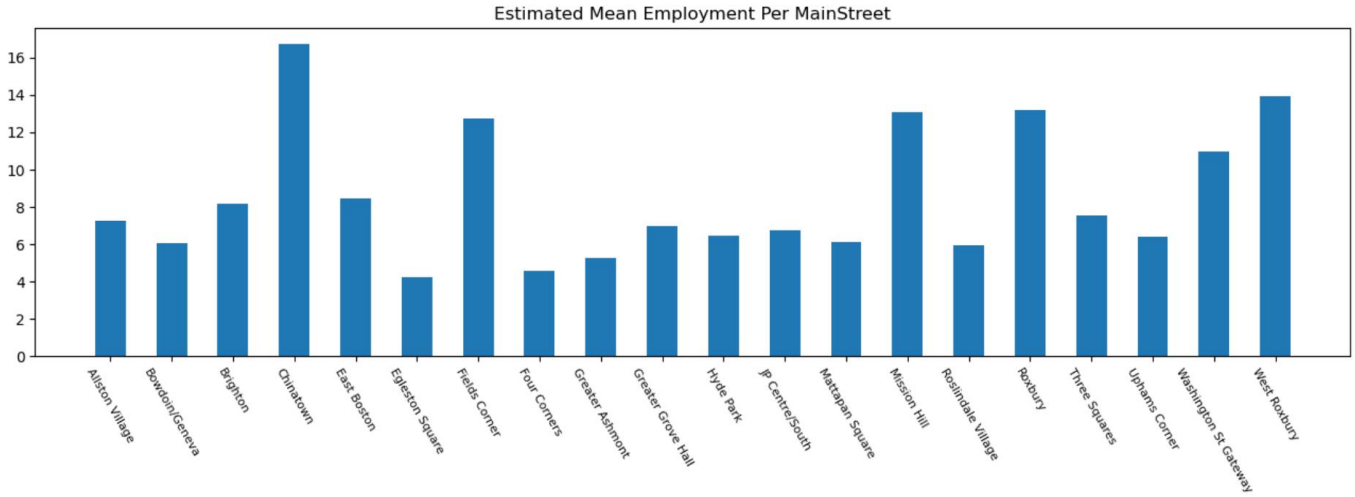
Bar chart shows the number of food businesses in each city. It is very easy to find that the number of food businesses in Boston far exceeds that of other regions. In other regions, the food business is more evenly distributed. From the pie chart which doesn't display here because of space, cities with 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.

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

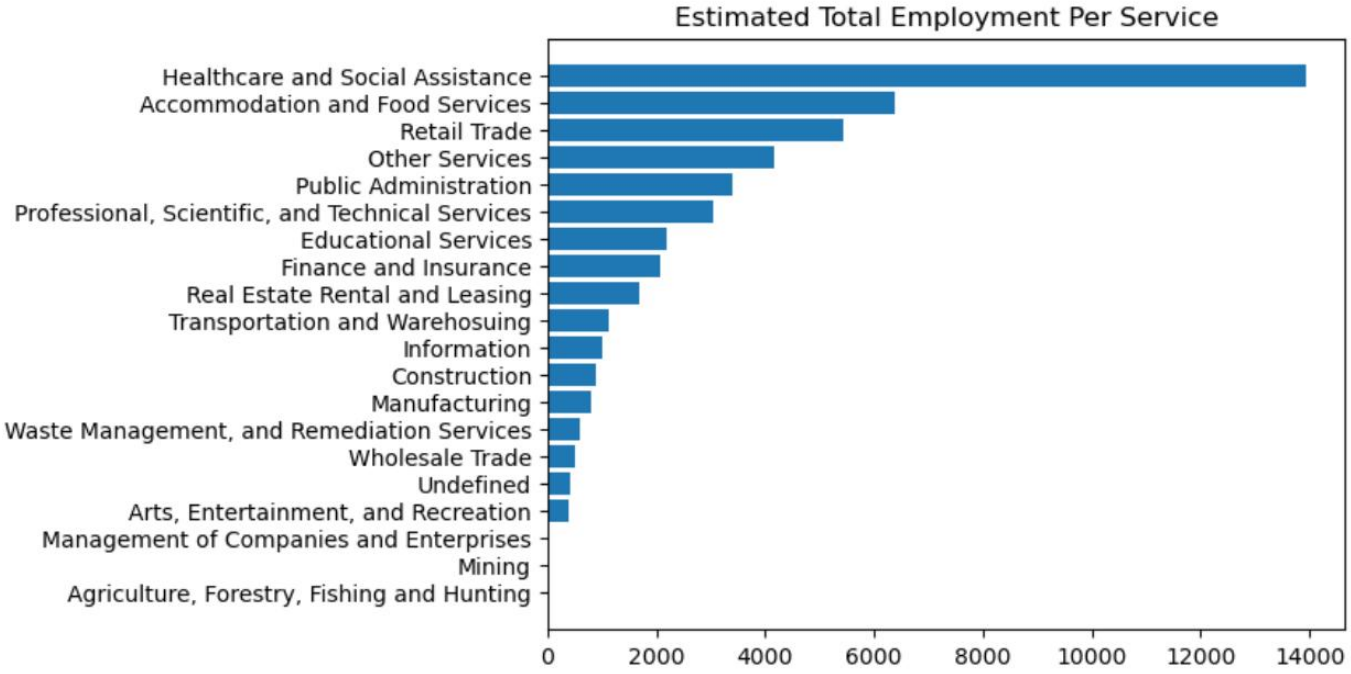
It's a new aggregated mobility dataset of individuals going to/from points of interest across businesses.

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.



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**

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, 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.

- **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.
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.

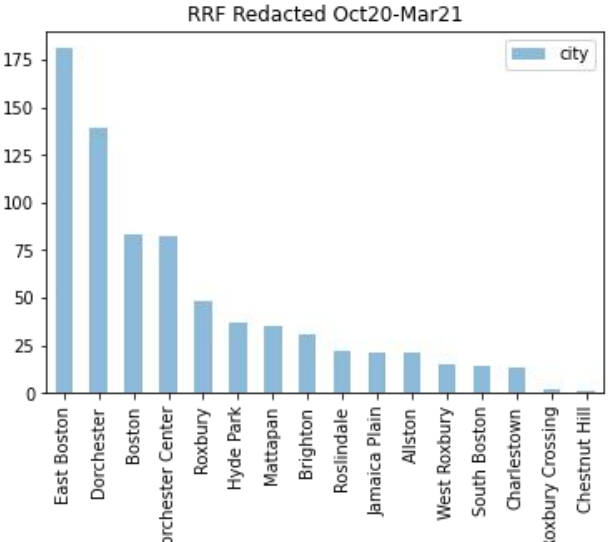
7. Suggestions for the future of the project

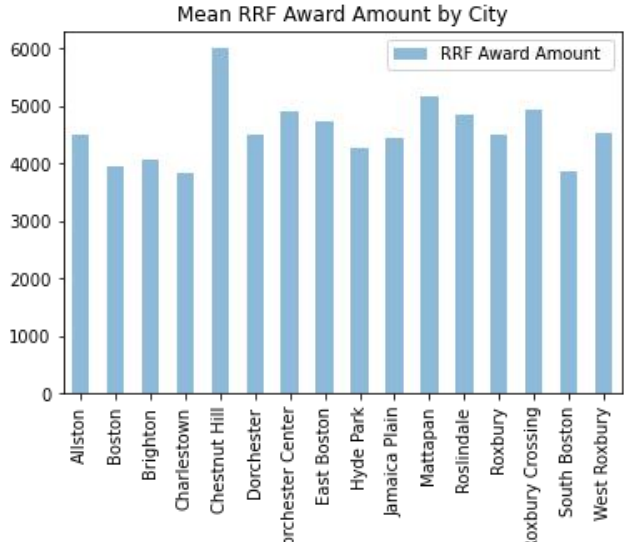
We need to further analyze the housing assistance data set, explore its connection with other datasets more deeply.

8. Extension project

The objective of the extension project is to further assess the rental and housing assistance data. In addition to evaluating if the rental relief funds were proportionally and adequately distributed across neighborhoods, examination of the pandemic's effect on data such as first time home ownership funds is desired.

For the extension project, evaluations to determine if areas with more rental relief fund allocations corresponded to fewer first-time home ownership fund allocations. Additionally, evaluations will be made to determine if areas with higher rental relief allocations also had more business relief applications. Intuitively, if residents need rental relief, they are less likely to spend money at restaurants and bars, which would correlate to businesses needing relief funds. By examining pre-pandemic data, determinations on whether area growth stagnated or declined during and after the pandemic.





Fall 2022 Councilor Mejia x City Services Project

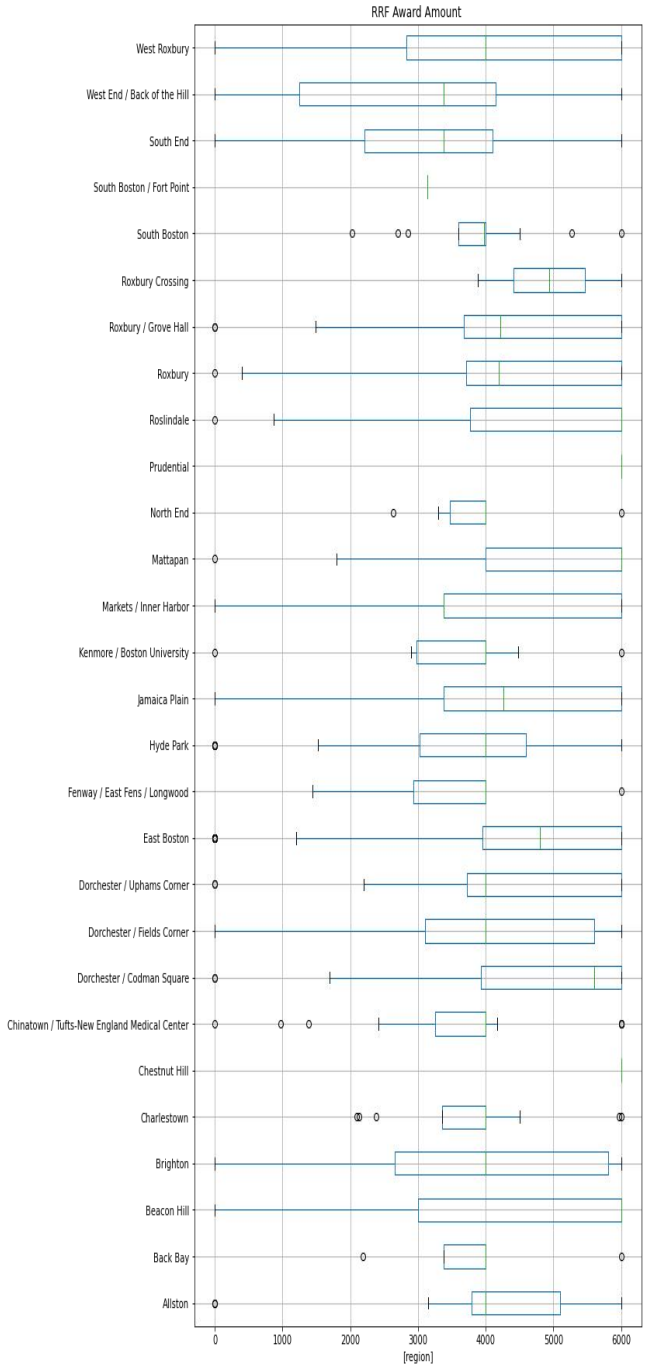
Team 3 - Draft Report

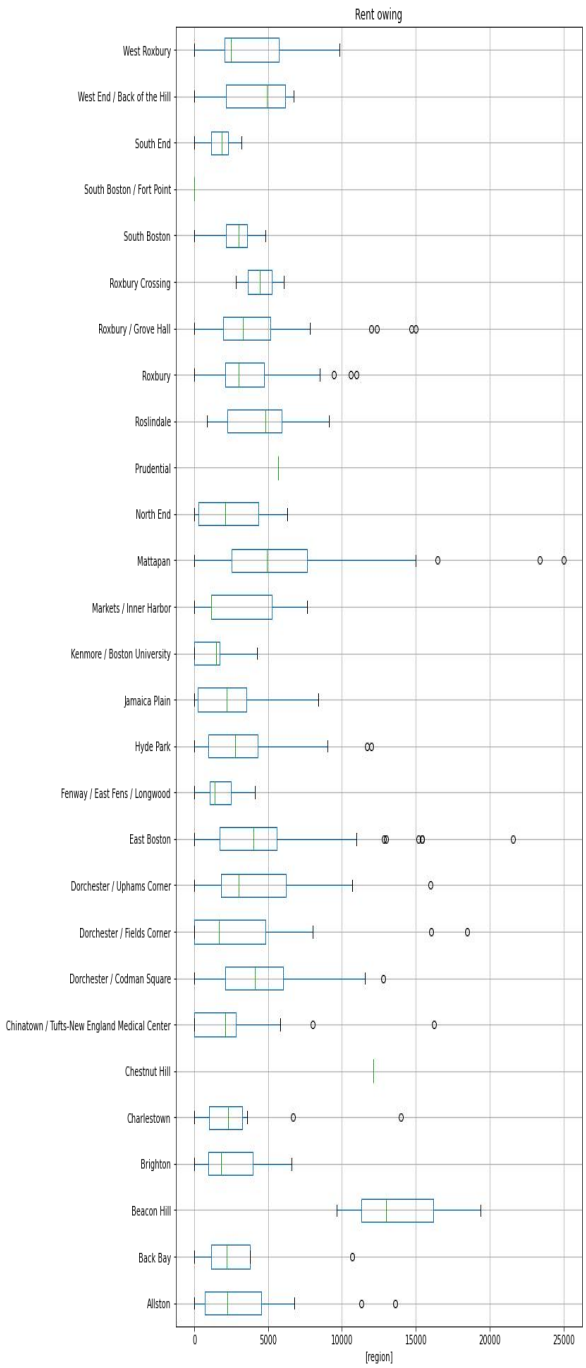
From the rental assistance dataset, we focused on the RRF Award Amount and Rent Owing. We plotted the box plot and bar chart for these two columns based on the neighborhood (zip code).

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.

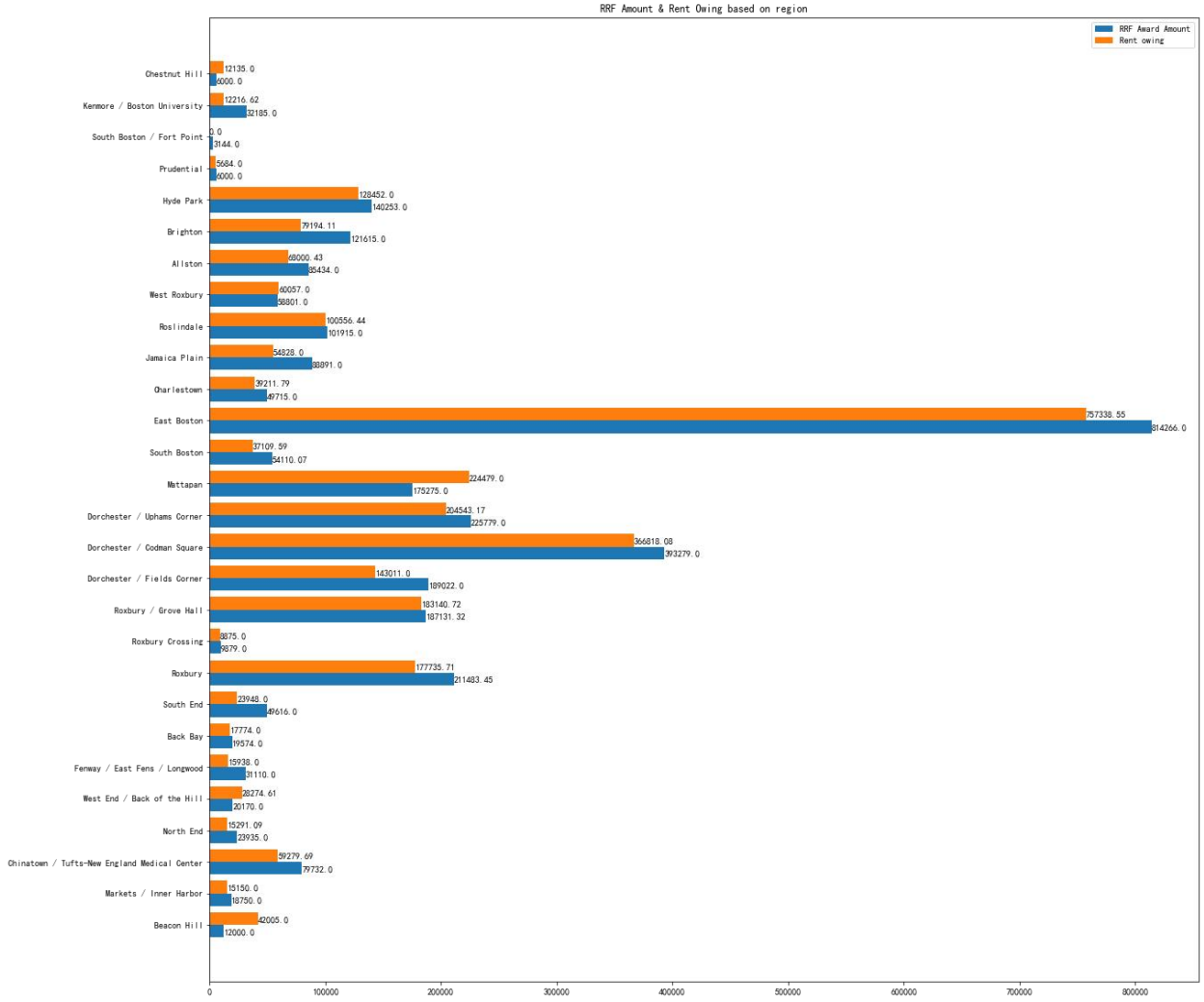
From the box plot about Rental owing, we can summarize that Beacon Hill had the highest amount of rent owing. And South Boston does not have rent owing.



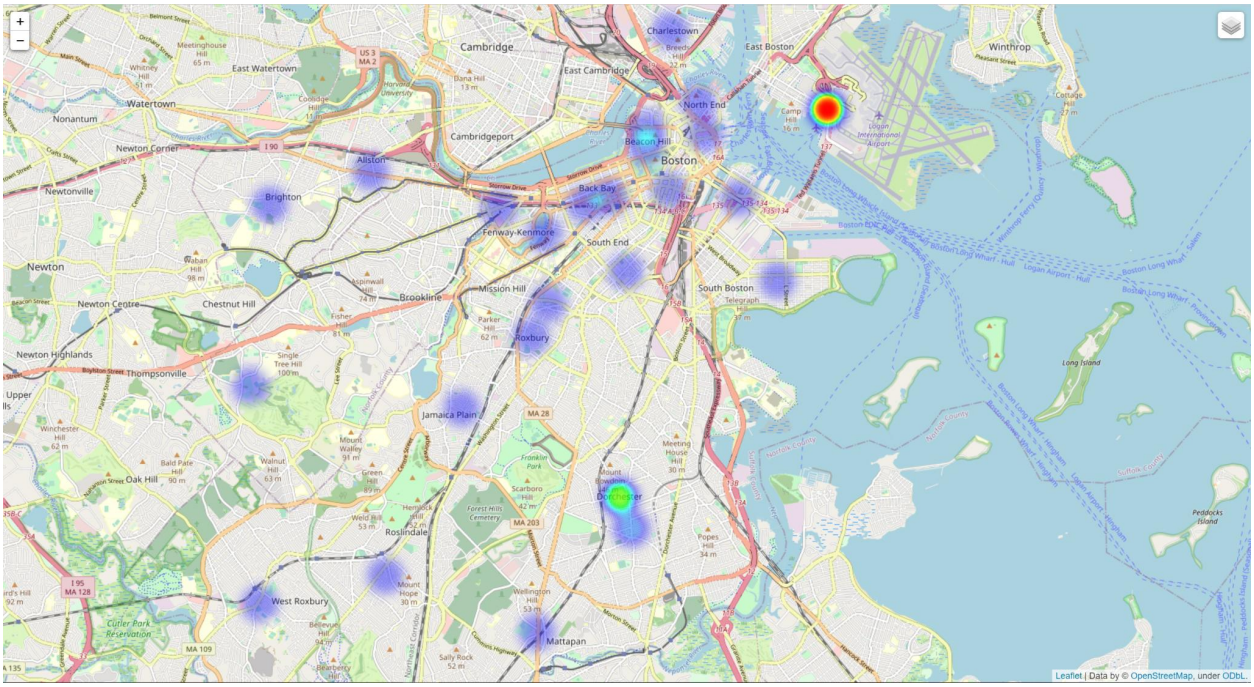


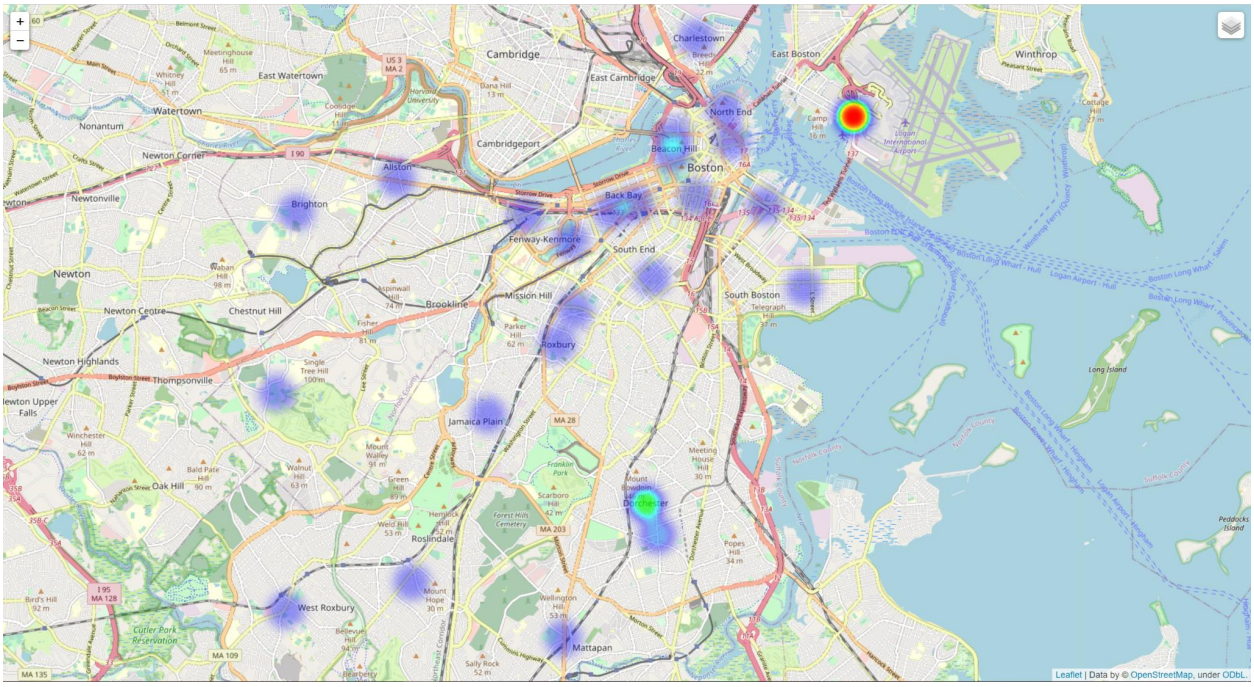
Fall 2022 Councilor Mejia x City Services Project

Team 3 - Draft 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.