

Final Report

Background + Motivation (what is this project about, who will use it, for what purpose, in what context etc.)

The goal of the project as a whole is to evaluate whether the city's funding is equitably distributed. A project of this scope advocates for positive change in the Boston community. What makes this project intriguing is that "equity" is the driving force behind the distribution. It ensures that the money is not being allocated to the "overarching community" but to those truly in need of the funds. We chose this project because it allows us to analyze the deciding factors that play a role in these distributions. It is important to analyze where the funds that we provide for the less fortunate in our neighborhoods actually end up. Through analysis of that data we can determine whether a redistribution of funds is necessary; e.g. we may need to add additional funds to a certain area or see if funds are being directed into the wrong hands. As members of the Boston community, it is necessary for us to be proactive and conscious of the factors that impact our environment.

Previous work (have there been other / different attempts toward these goals in the past?)

For this project specifically, money was already being distributed amongst communities in previous years. Our goal was to ensure that the money is being equitably distributed. To the best of our knowledge, this is a new undertaking and in-depth research has not been conducted on this in the past.

Data Collection Method

We analyzed the census data to determine the communities that we should expect to be targeted by the city's programs and align these findings with our PM. To achieve an equitable distribution of funds, communities of color and poorer communities should be receiving a larger proportion of funds and licenses than wealthier white communities. We determined that Dorchester, Roxbury, Mattapan, and Hyde Park are the neighborhoods that should receive the largest proportion on funds.

We then analyzed the data on business and liquor licenses administered by the city of Boston. The result showed that Dorchester is the largest receiver of these licenses.

The analysis of the Business Grants dataset also showed Dorchester as the largest receiver of aid. However, by normalizing this data according to population, we can observe that Back Bay and Chinatown/Financial District were actually the largest recipients of the business grants(per capita) throughout the Covid pandemic.

We performed analysis on the Capital Investment dataset for FY21-25 that represents approved and ongoing projects in the city of Boston, including details on the department that is executing the project and the neighborhood it is occurring in. From this data, we can observe behavior of each department in the city of Boston. For example, the Parks and Recreation department has the largest number of ongoing projects, but each project has a small budget compared to the project budgets of other departments. Furthermore, the Parks and Recreation department allocates a smaller total amount of money to Capital Investments than other departments.

We can see a change in this data from our previously analyzed datasets, as Dorchester does not receive a large number of Capital Investments. In fact, if we normalize this data according to population, we can observe that the Harbor Islands is the largest recipient of Capital Investments by far. Some investigation into these projects shows that many of the approved projects are large renovations to the islands and the bridges connecting them and a planned renovation of the Fire Department Academy located on these islands. This “neighborhood” has a population of 401 people according to the census, so these large projects skewed the normalization significantly. By removing this as an outlier, we can still observe that the target communities are not receiving a significant amount of capital investments compared to neighborhoods like Charlestown.

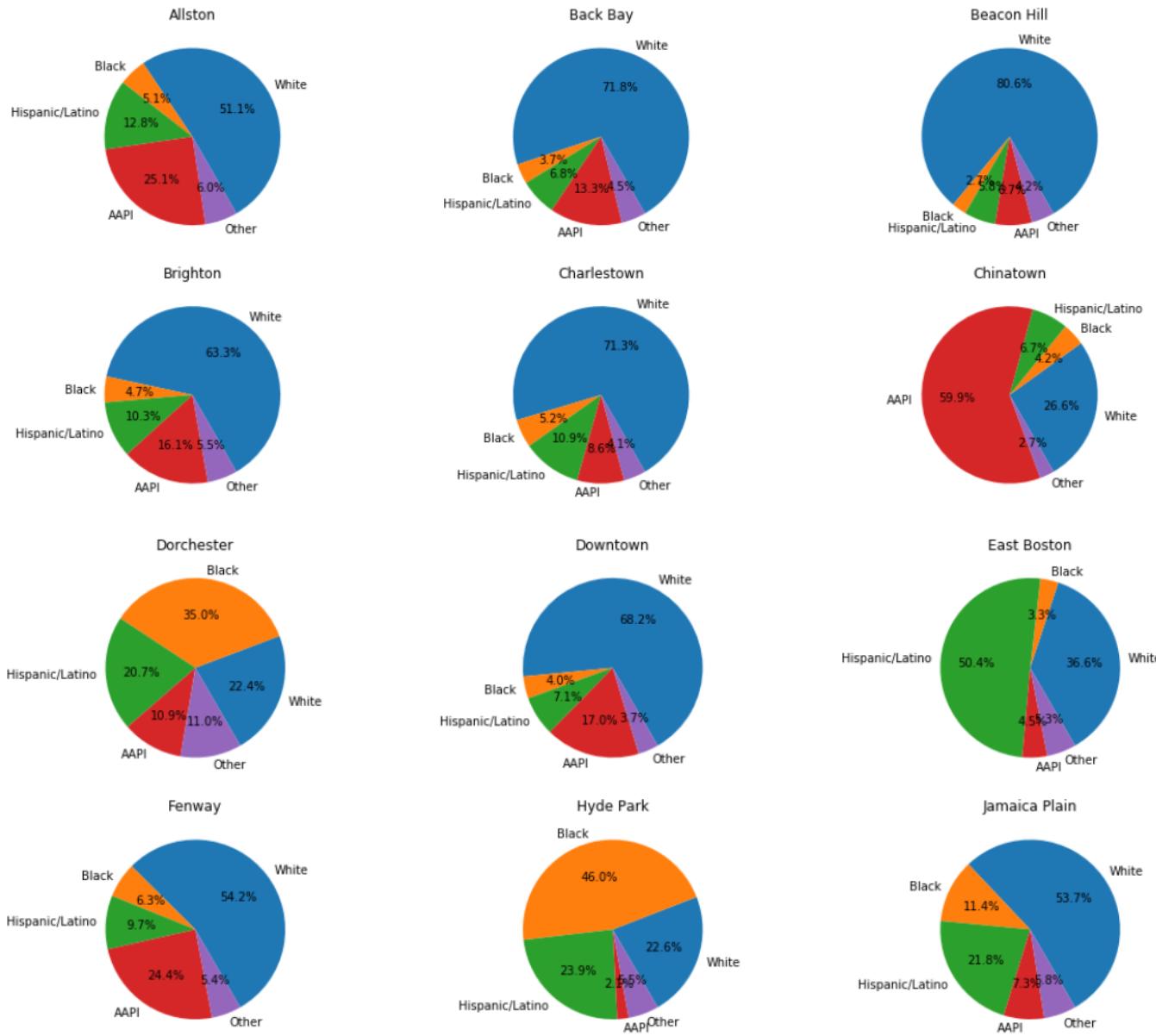
We performed analysis on the Rental Assistance data to conclude that the amount of money being distributed to citizens to support their needs during the pandemic has been equitably distributed according to our target neighborhoods.

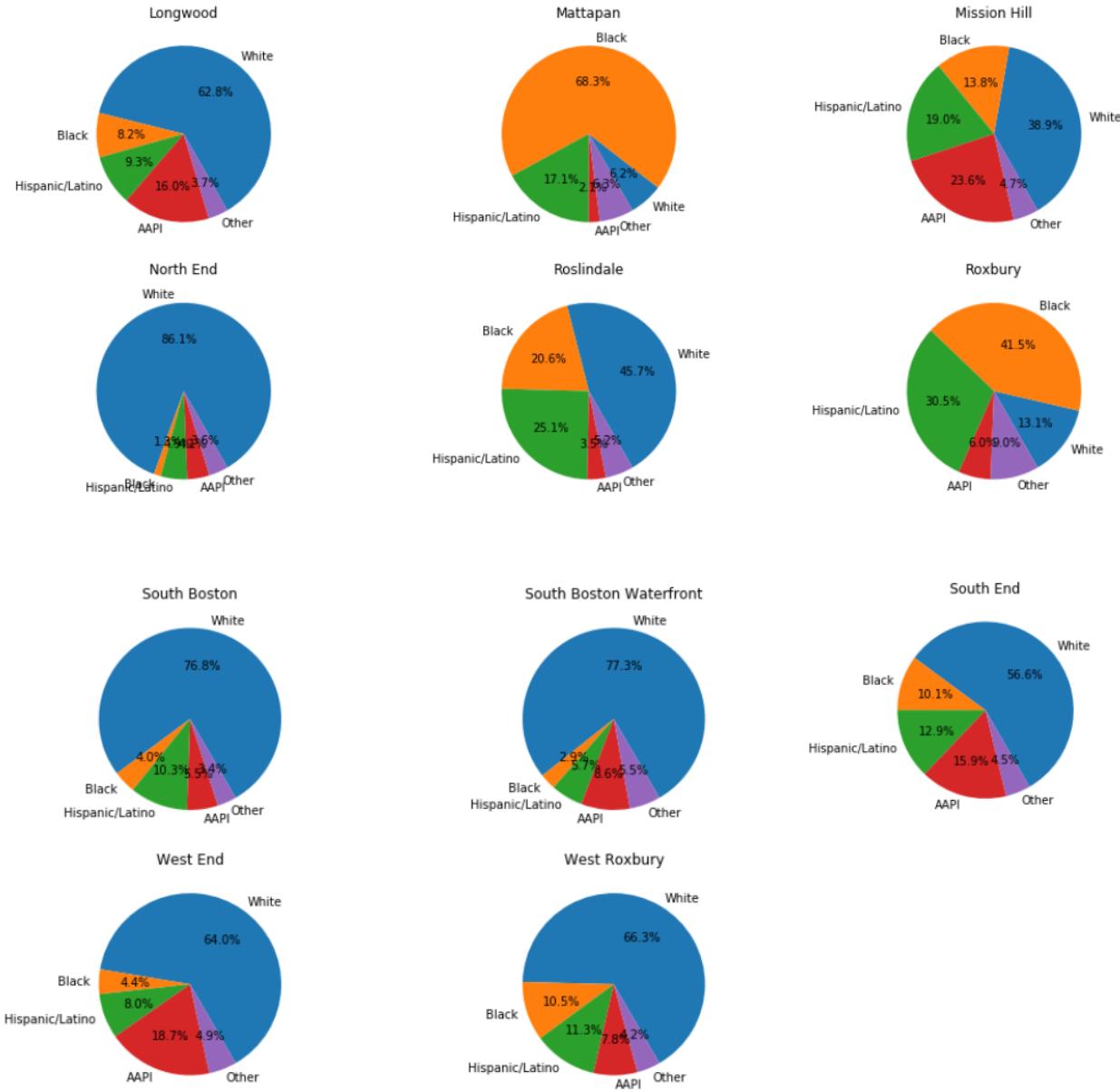
Finally, for our extension project, we performed analysis on data about service calls(311) to each department throughout the city of Boston. We determined which neighborhoods had the greatest number and proportion of overdue service requests and which departments were responsible for those overdue requests.

Data Visualization and Exploration:

Census Data:

We analyzed the racial demographics of each Boston neighborhood using census data(alphabetical order):

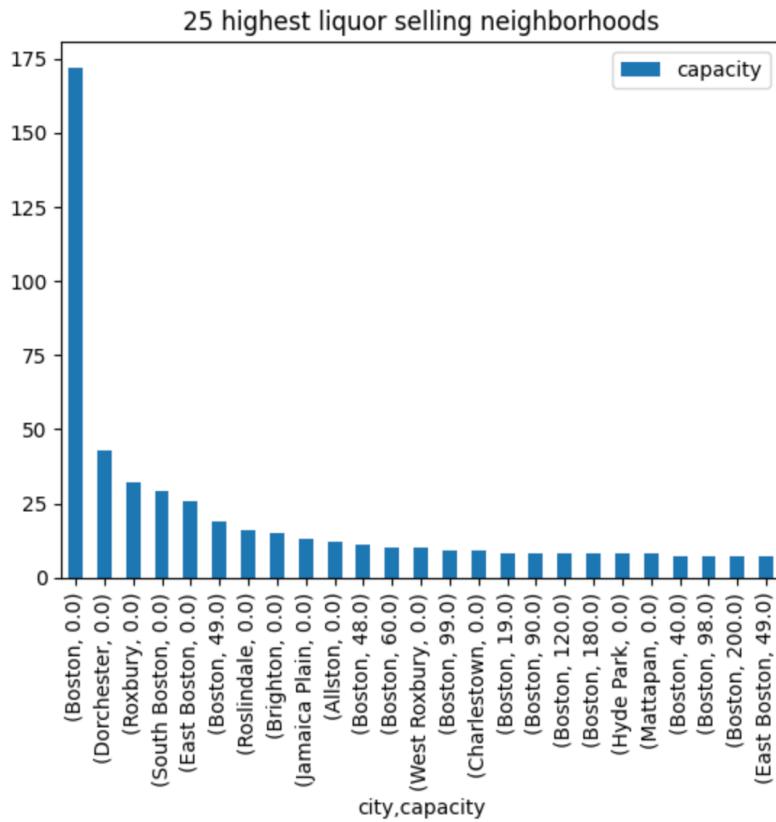




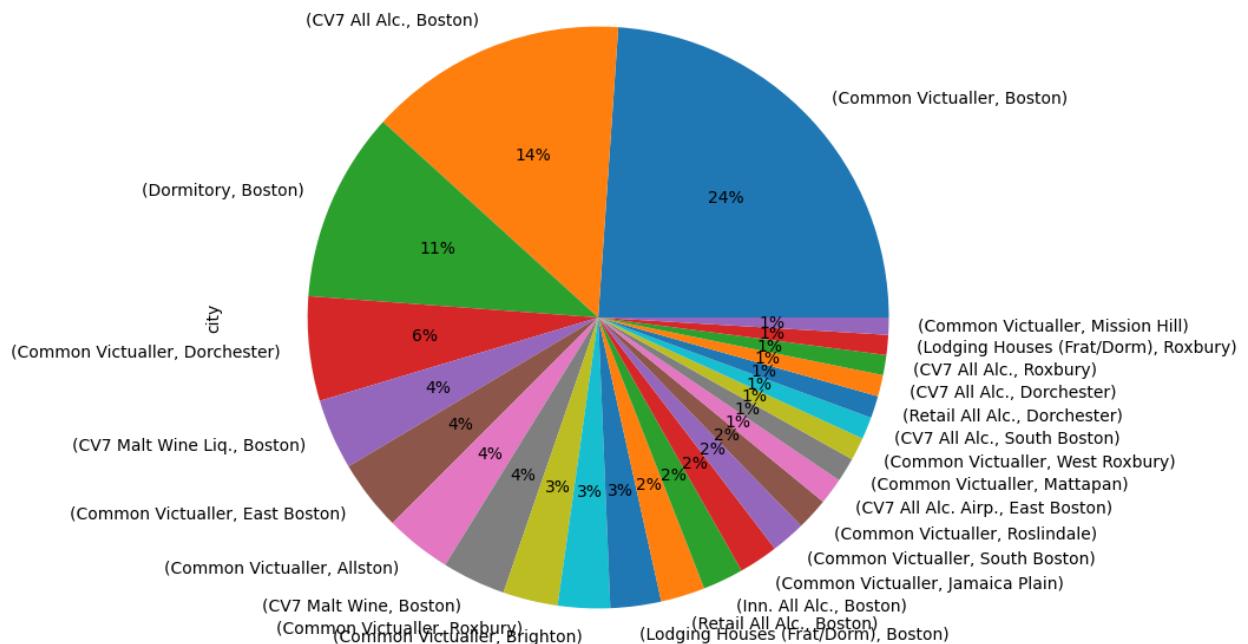
From these visualizations, we can identify the target neighborhoods based on demographics to be Dorchester, Mattapan, Roxbury, Hyde Park, East Boston, and Chinatown since these neighborhoods are not majority white.

Liquor Licenses:

We analyzed data about the number of liquor licenses administered to business in Boston neighborhoods.



Percentage and type of licenses sold in each city in the Boston Area

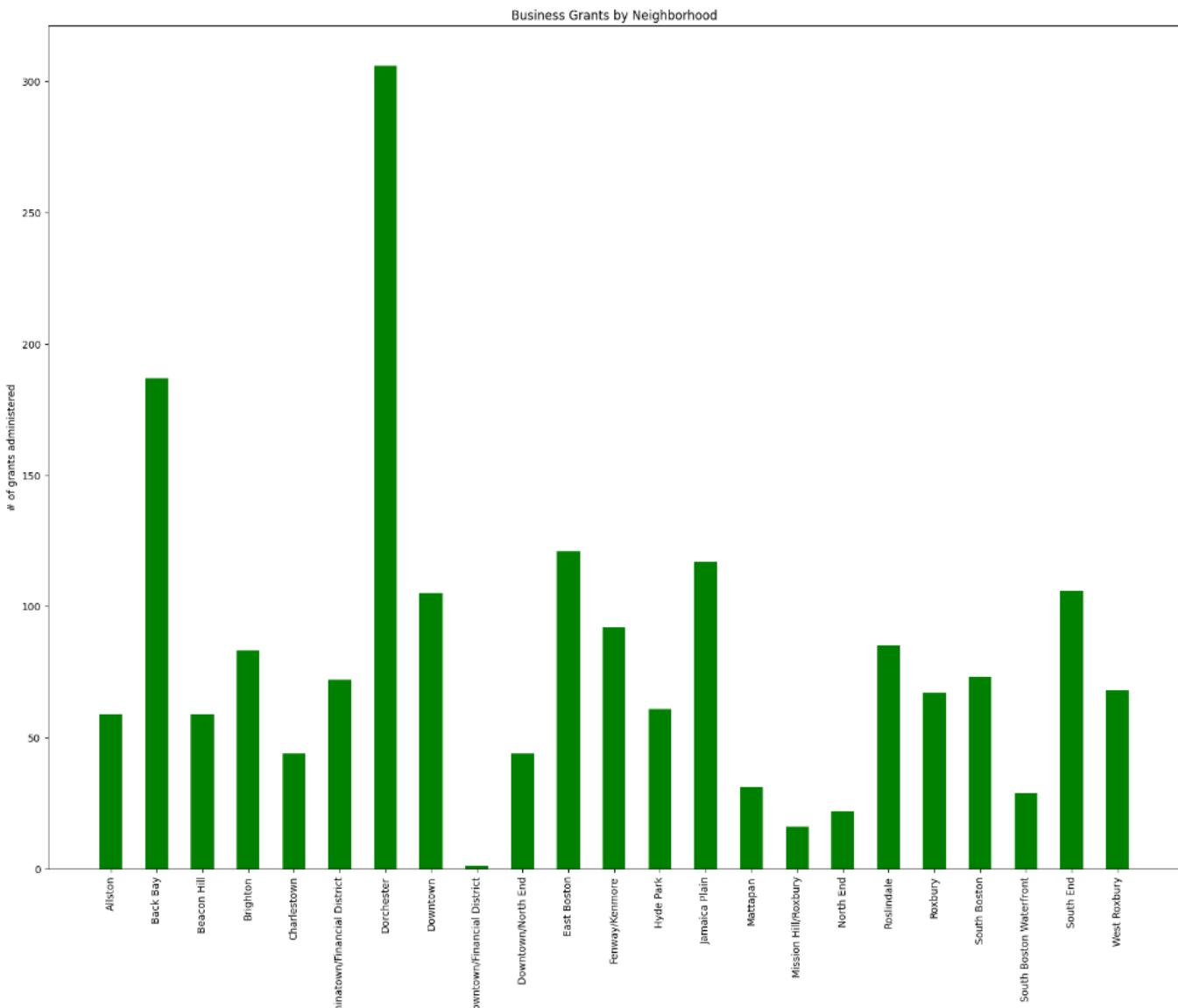


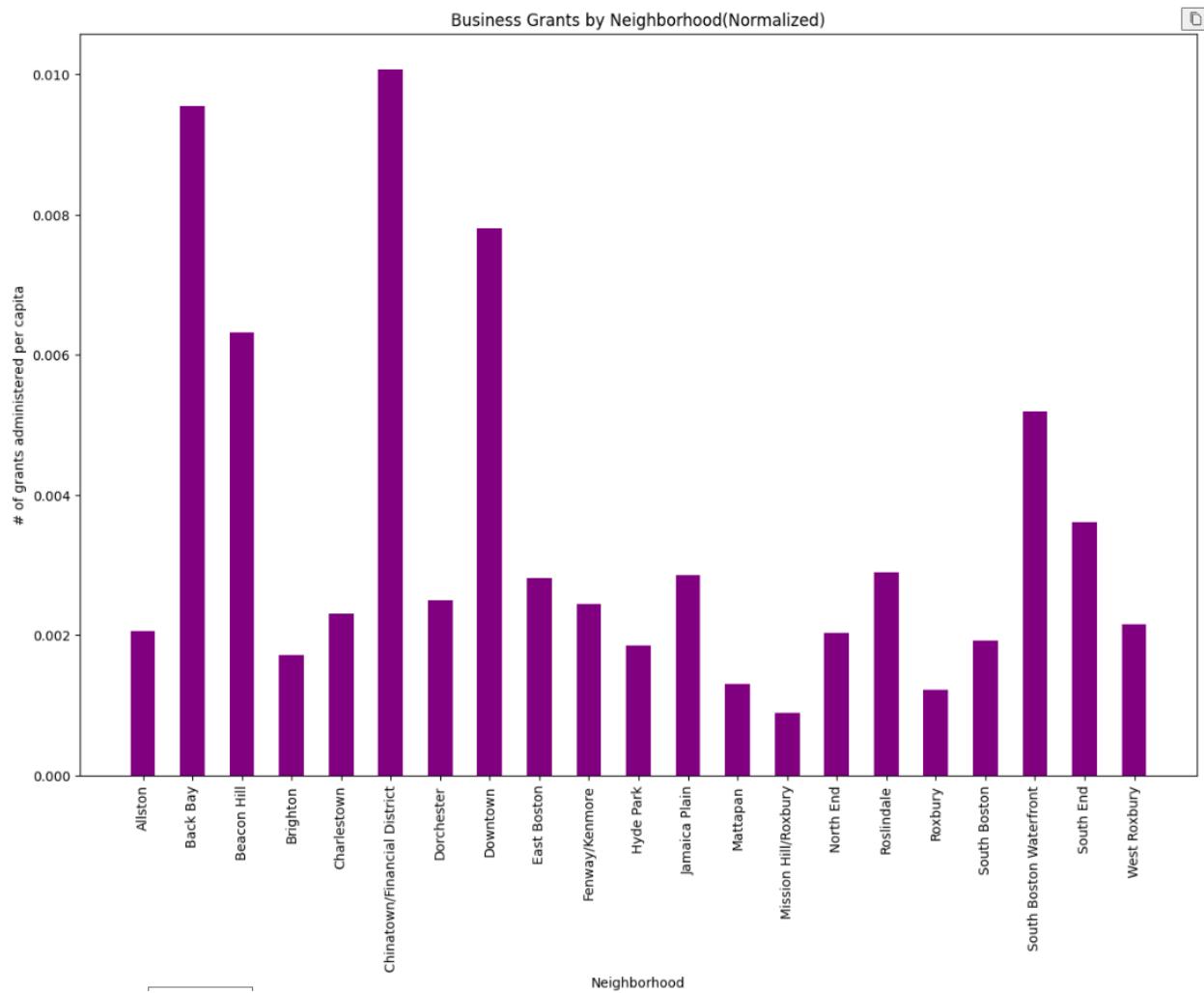
Liquor and Common Victualler(license to cook and sell food) licenses support these neighborhoods by providing them with the resources for businesses to thrive. These licenses are integral in the maintenance of a communities' economy, and it is important that local businesses have the resources to be successful. We can observe from this analysis that Dorchester is the largest recipient of these liquor and common victualler licenses. Because Dorchester is a large neighborhood with racially and economically distinct sub-neighborhoods, a further analysis would look into what areas of Dorchester are receiving the most licenses. Furthermore, this analysis reveals a limitation of the data that we analyzed. Many licenses are listed as being administered to Boston, rather than a specific neighborhood. This makes it difficult to determine whether these licenses are being equitably distributed across the Boston neighborhoods. Much of the data below has the same limitation.

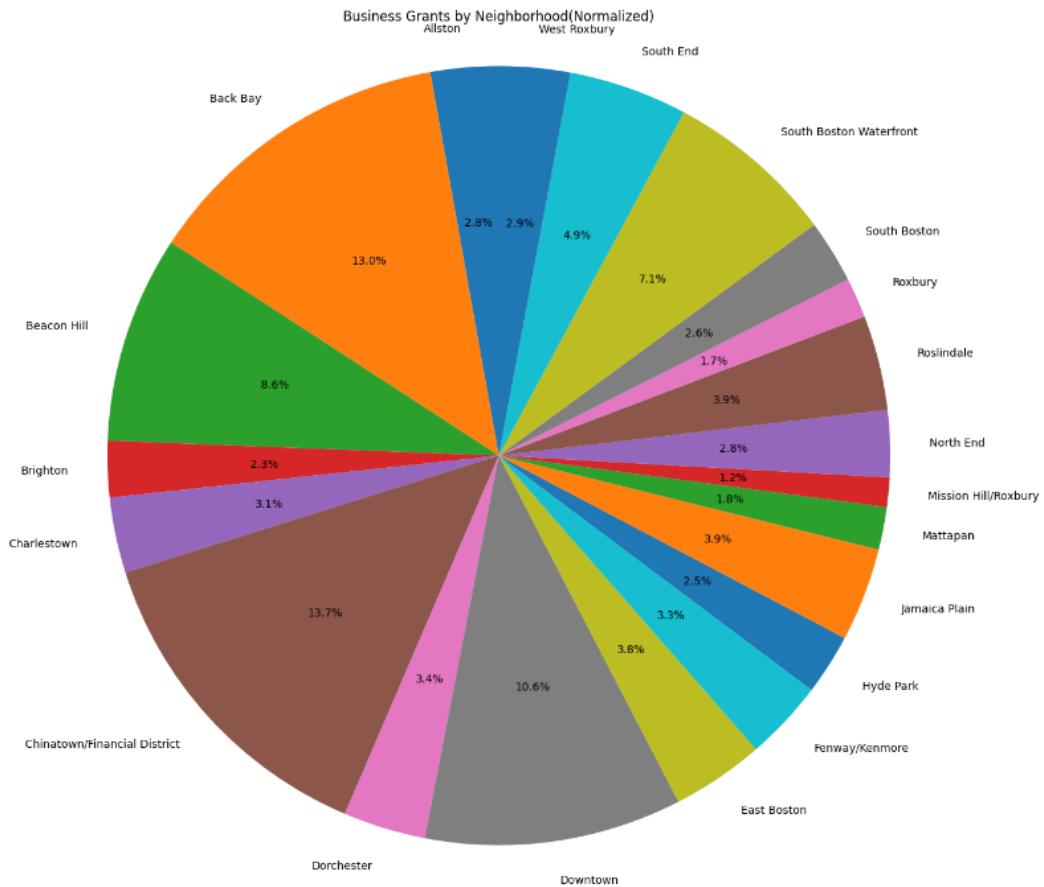
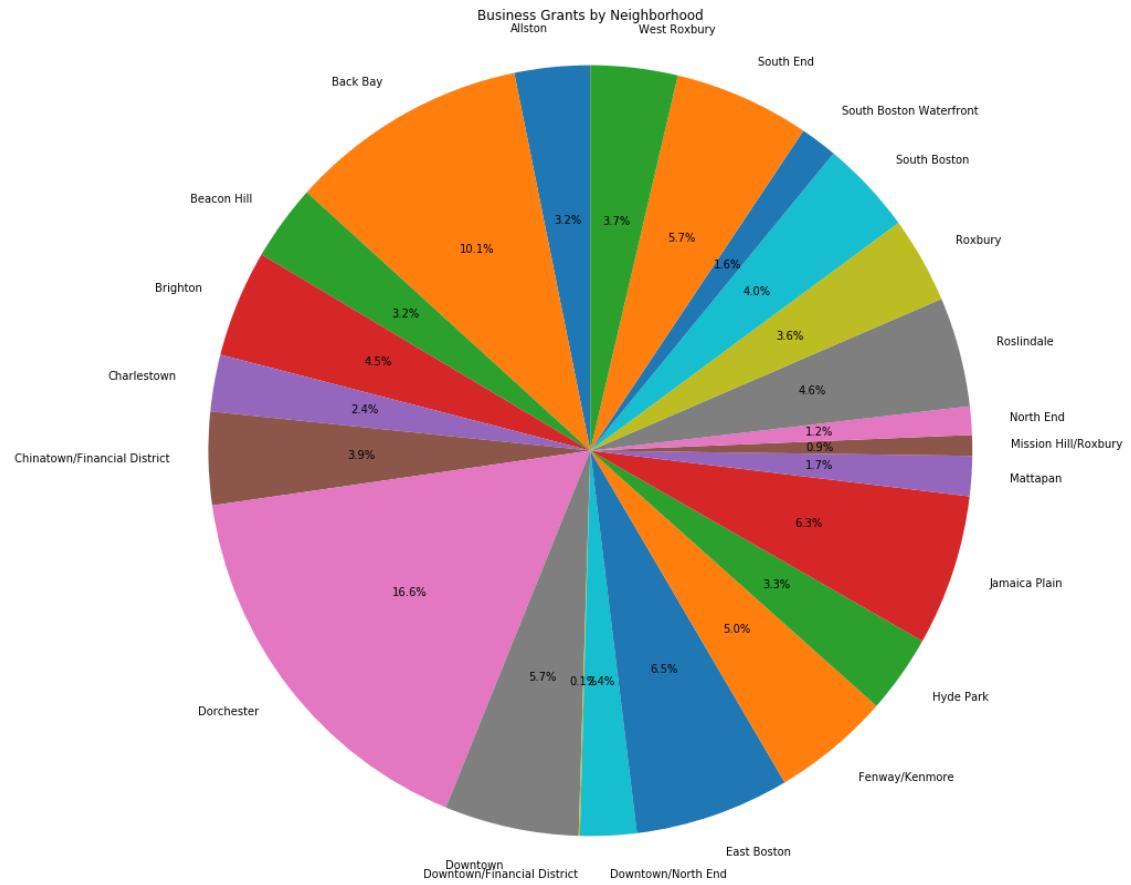
Business Grants:

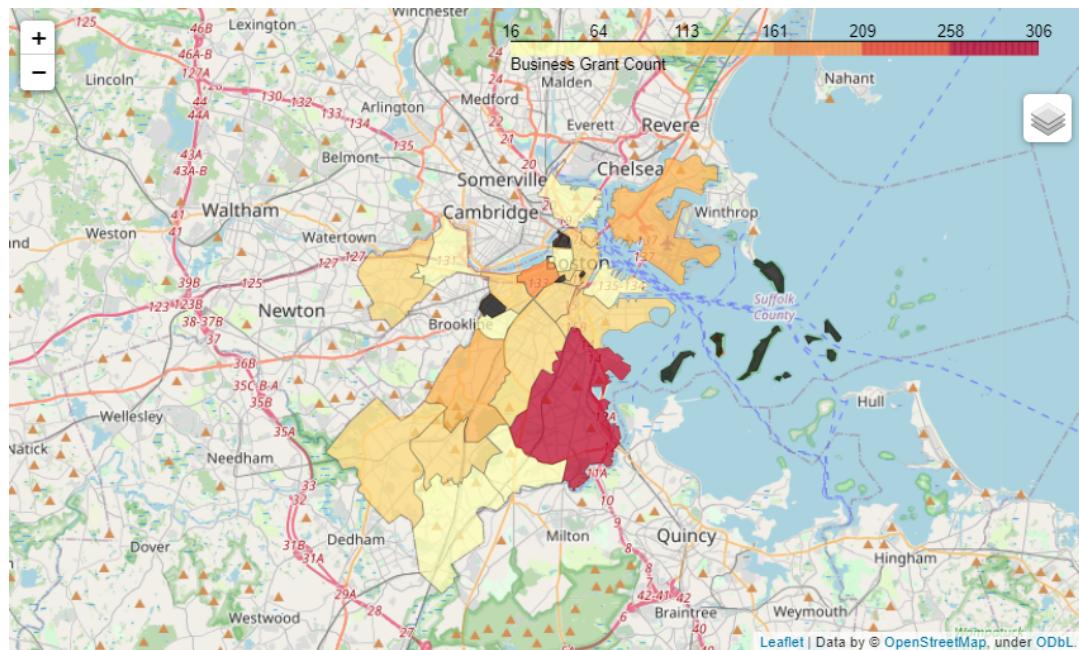
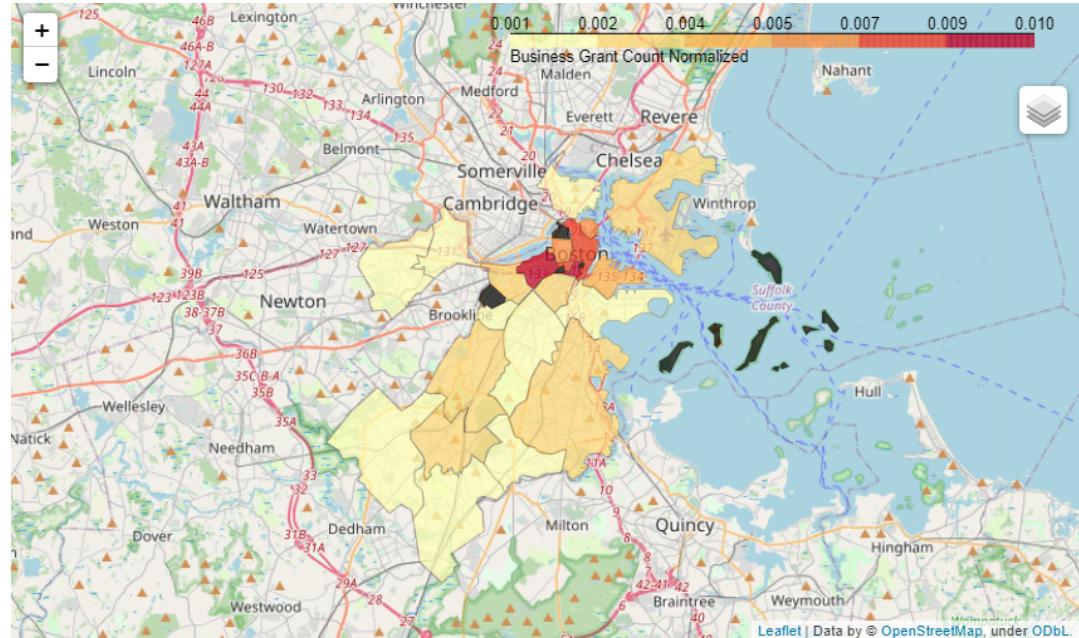
We analyzed data about where business grants were administered during the pandemic. These visualizations show the data before and after it being normalized by population.

Original:



Normalized:



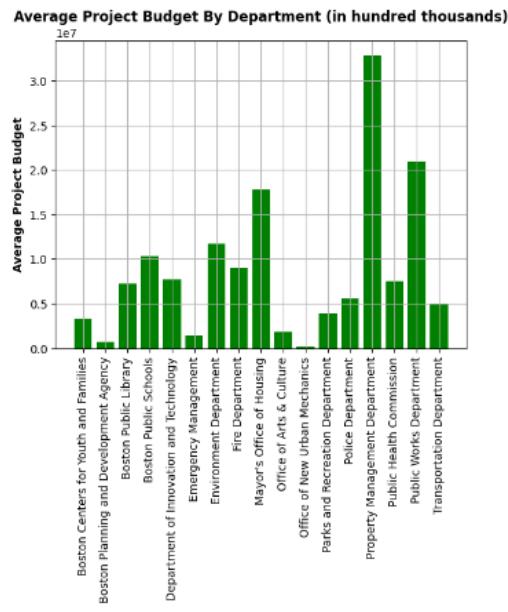
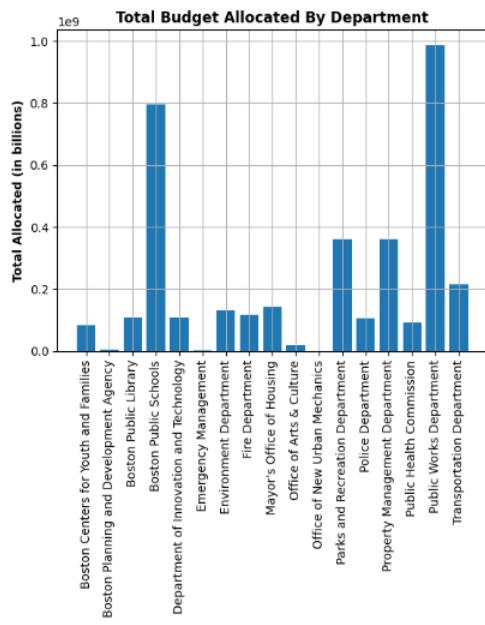
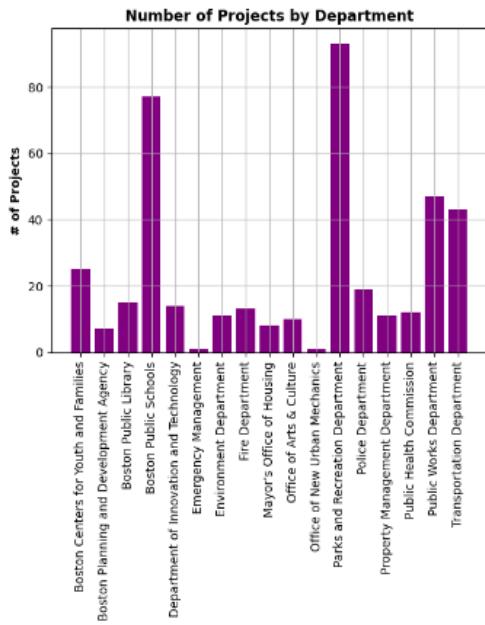
Original:**Normalized:**

Similar to the analysis of the liquor licenses, this data is significant because it informs us which neighborhoods were the recipients of the most support and resources to succeed throughout the pandemic. We can first observe that Dorchester was the recipient of the highest number of business grants, as 306 businesses received grants. Chinatown and Back Bay, however, received the highest number of grants per capita, .01 and .009 respectively. In comparison, the neighborhoods that we identified as “target neighborhoods” to receive business grants received the following:

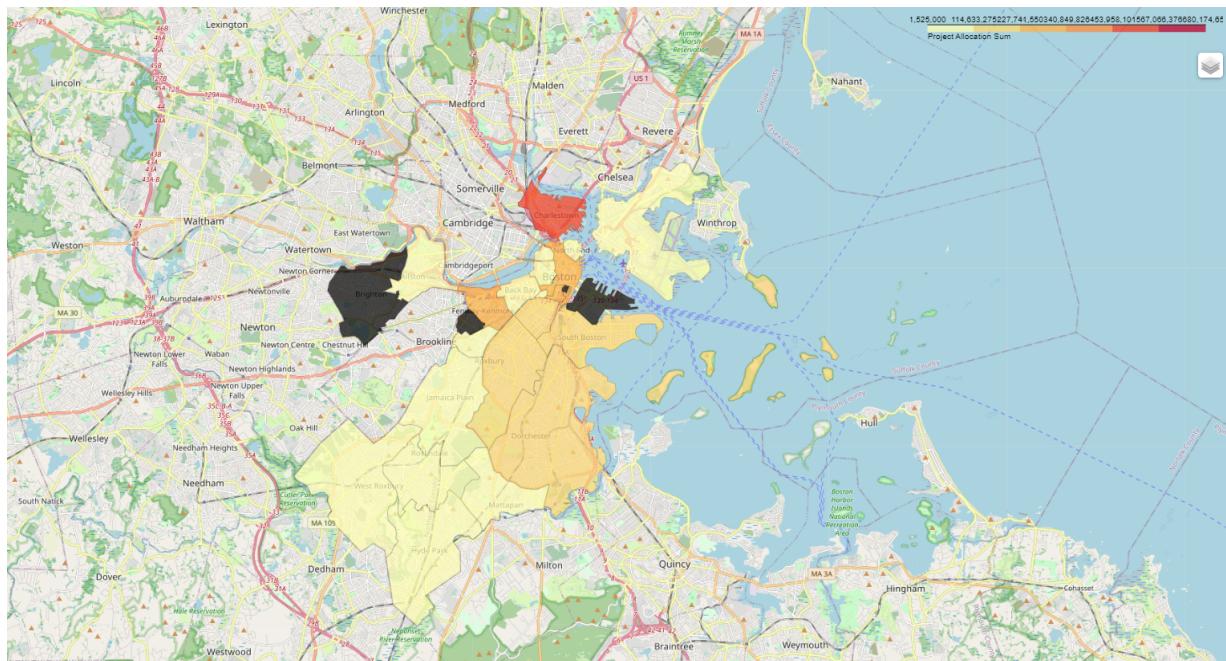
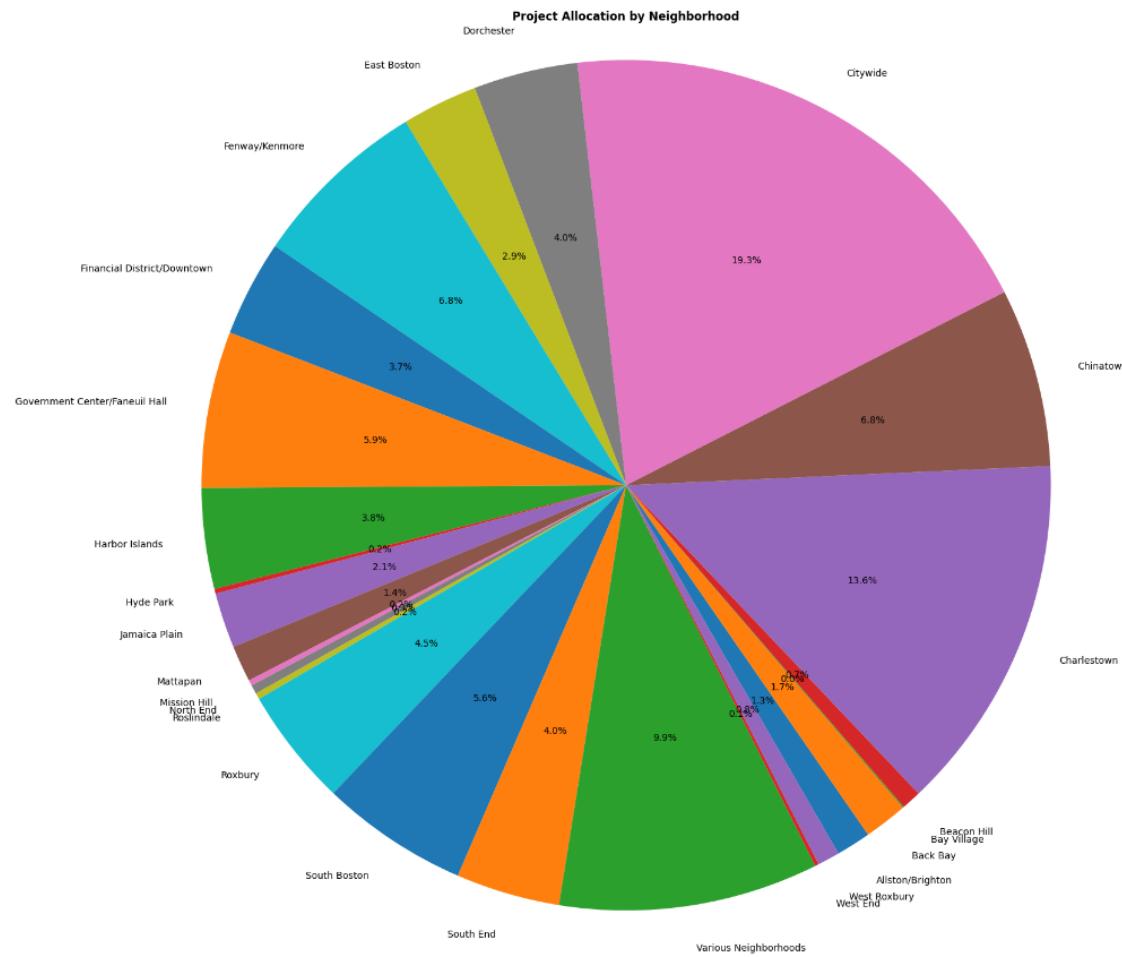
Dorchester: 306 business grants, .0025/capita
 Mattapan: 31 business grants, .0013/capita
 Roxbury: 67 business grants, .0012/capita
 Hyde Park: 61 business grants, .0013/capita
 East Boston: 121 business grants, .0028/capita

Capital Investment:

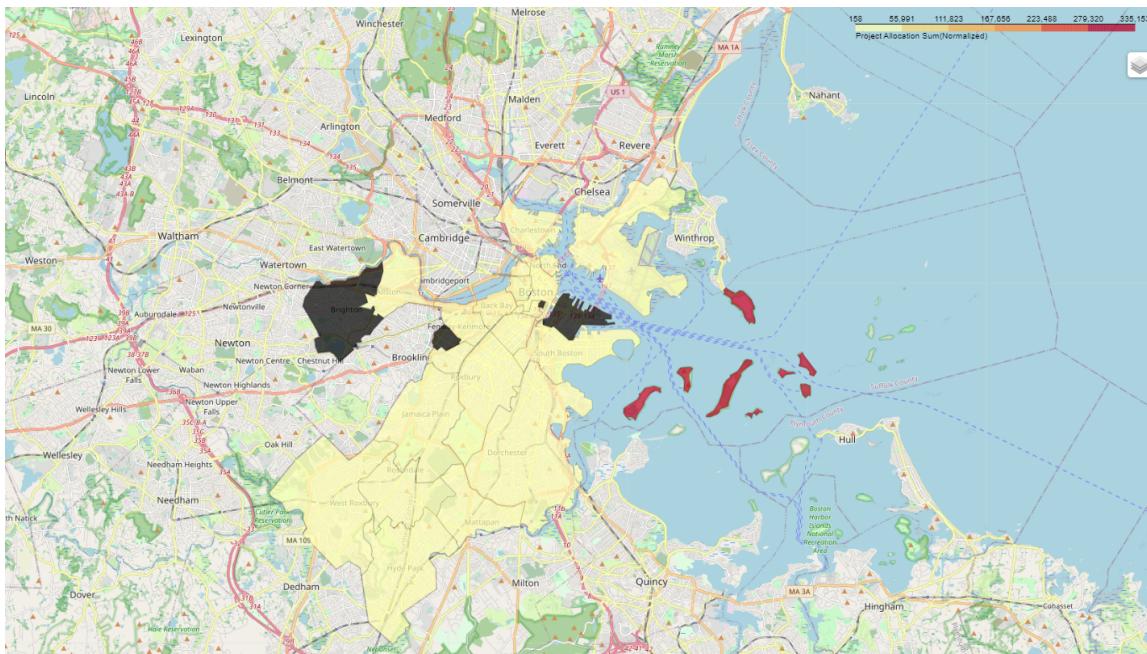
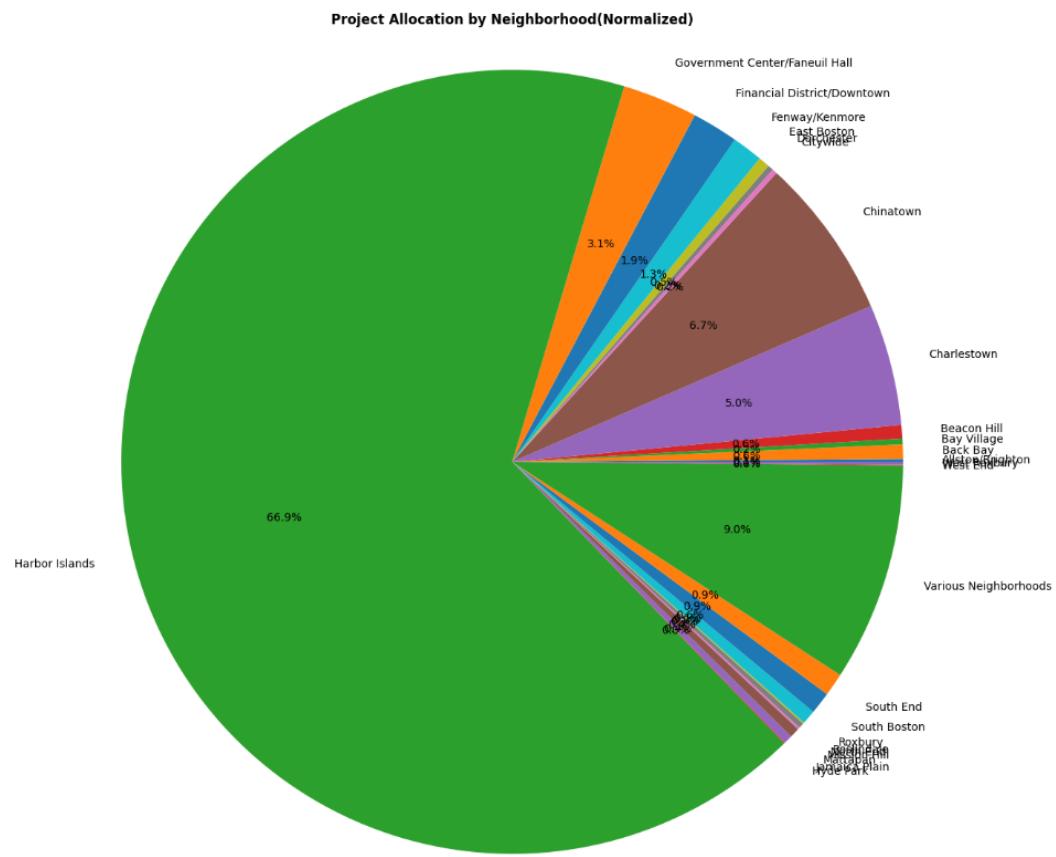
We analyzed the city of Boston's Capital Budget Plan for FY2021-2025.



These visualizations are helpful in understanding the behavior of each of the city's departments. For example, we can observe that the Parks Department has many projected projects for this fiscal cycle, but each project has a relatively low budget compared to the projects of other departments. We can observe that the Public Works Department has the highest budget to allocate, with a budget of nearly \$1 Billion.

Original:

Normalized:



These visualizations inform us of the proportion of the city's total capital budget each neighborhood is projected to receive in the upcoming fiscal cycle. Similar to the liquor license

data, we can observe that many projects are projected to occur in “various neighborhoods” or ‘Citywide’. This reveals a limitation of our conclusions because we cannot determine the equitability of funds allocated to “various neighborhoods”.

Normalizing the data identifies Harbor Islands as the largest recipient of capital investments per capita. These islands have a population of ~400, which greatly influences the normalization. The Fire Department Training Academy is located on the Harbor Islands, which has large renovations planned for this fiscal year, which accounts for the large amount of money they are receiving.

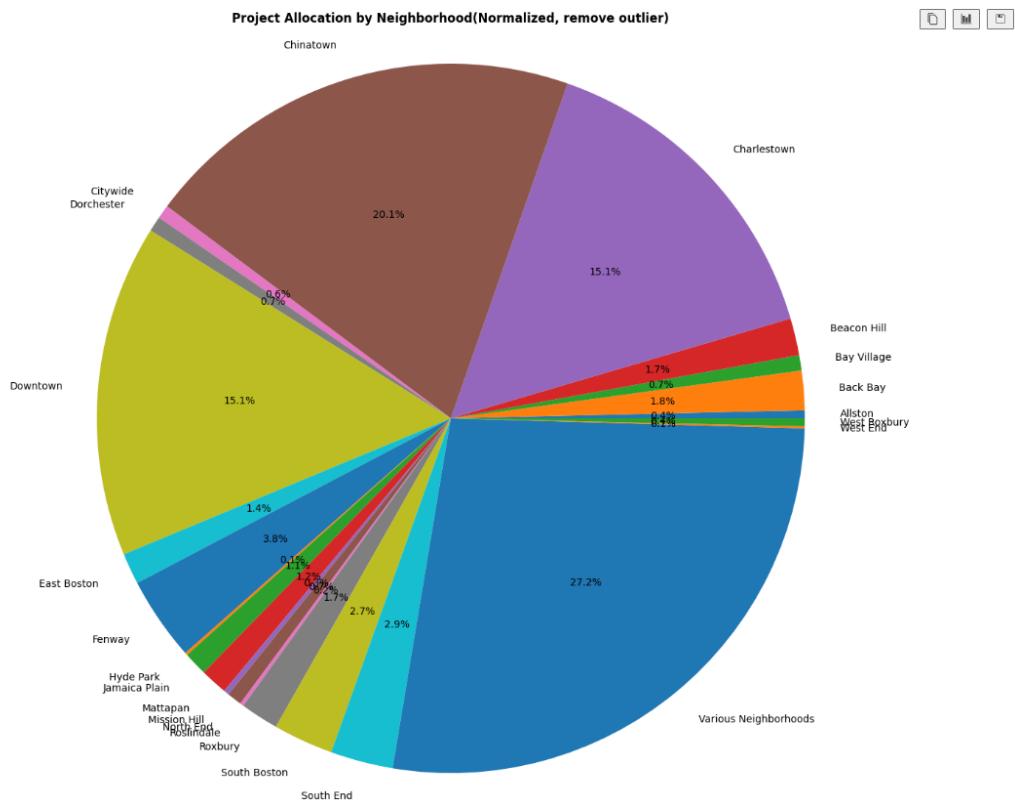
Because the amount allocated to the Harbor Islands is more than two standard deviations above the average capital investments of a Boston neighborhood, there is the option to remove it as an outlier of this data.

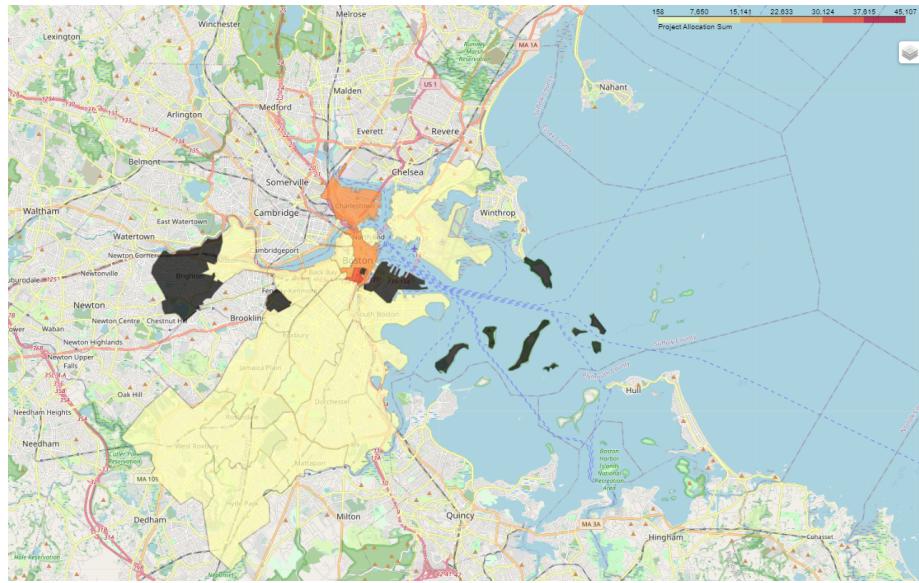
Average allocation = \$1.466231e+08

Standard Dev of allocation = \$1.710563e+08

Harbor Islands Allocation = \$6.801747e+08

The following visualizations are the result of removing the Harbor Islands as an outlier:



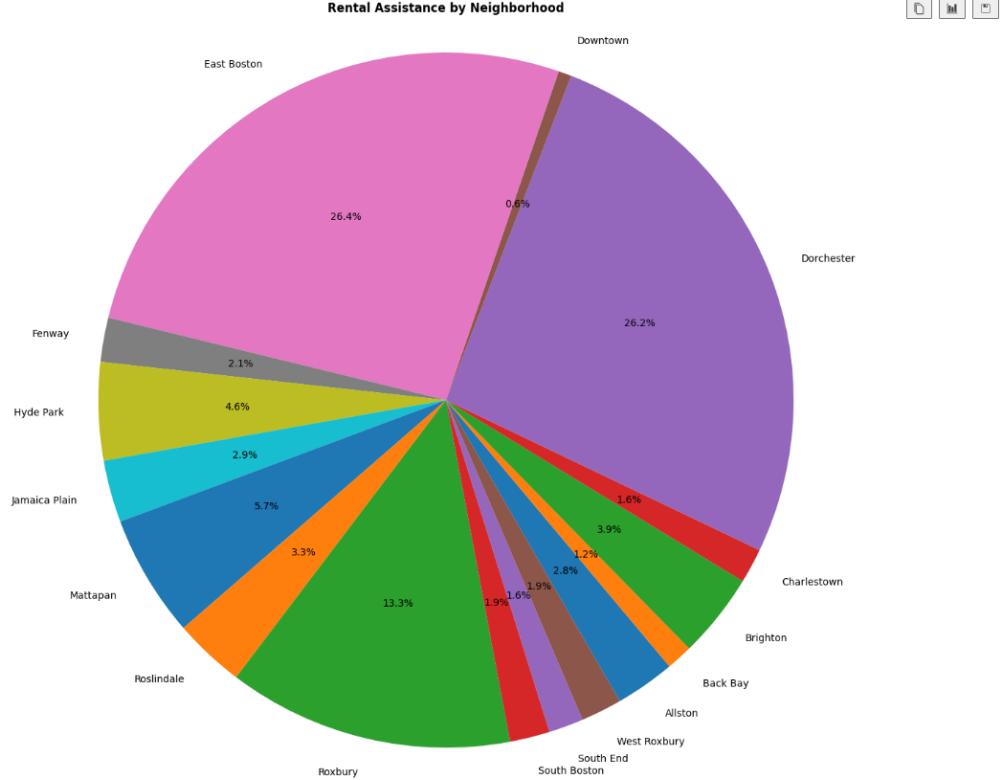
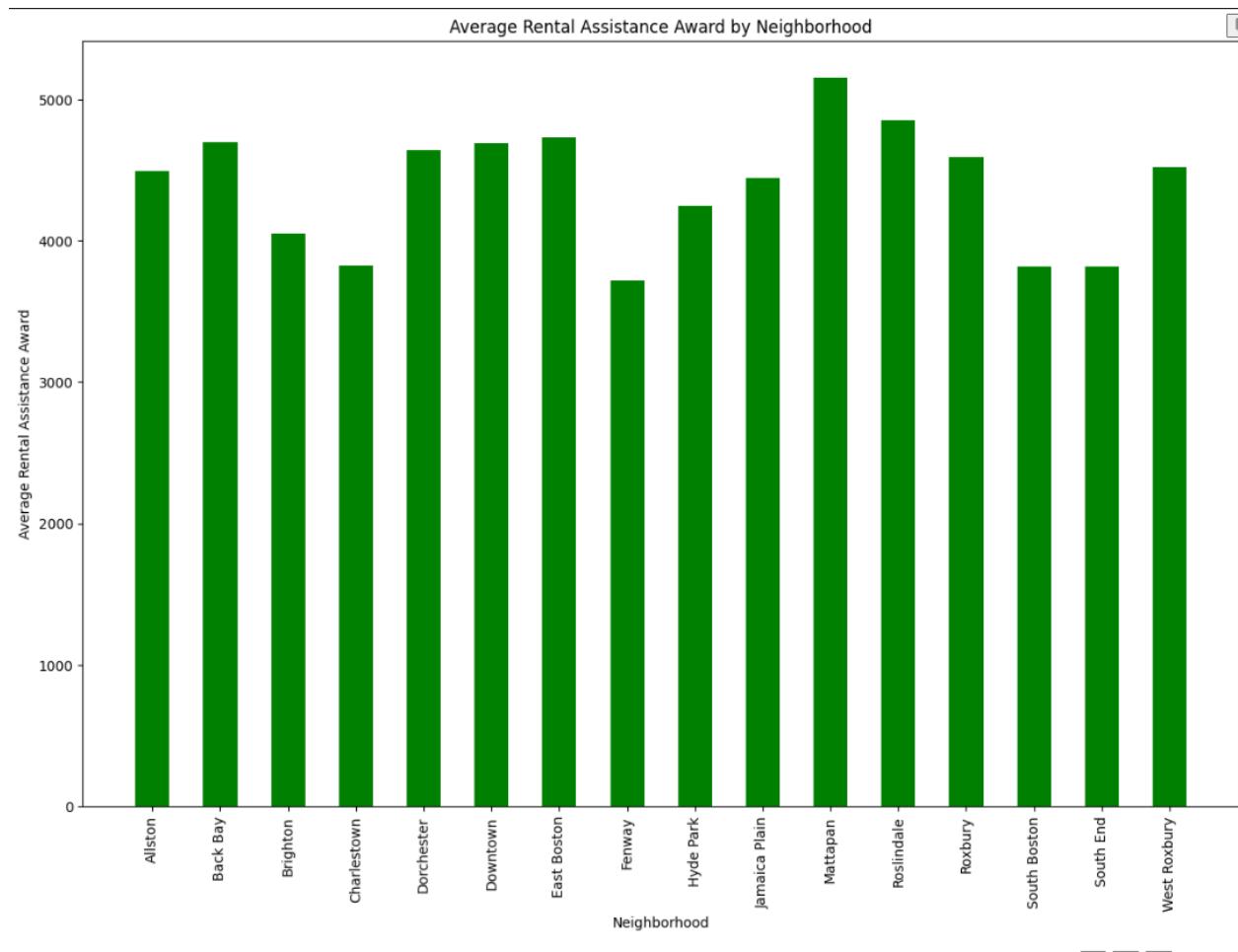


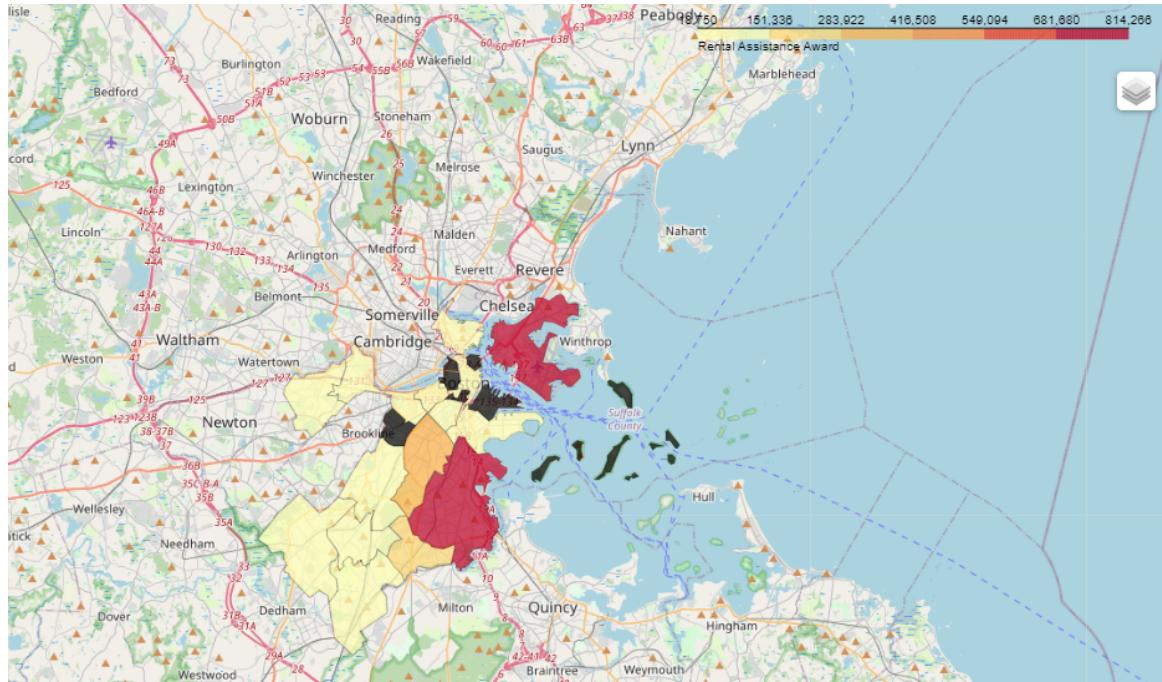
With or without removing the outlier, we can observe that our target neighborhoods will not be receiving as much money in capital investments in the upcoming fiscal cycle as Chinatown and Charlestown, both of which are wealthy neighborhoods in comparison to our target neighborhoods.

Rental Assistance Funds:

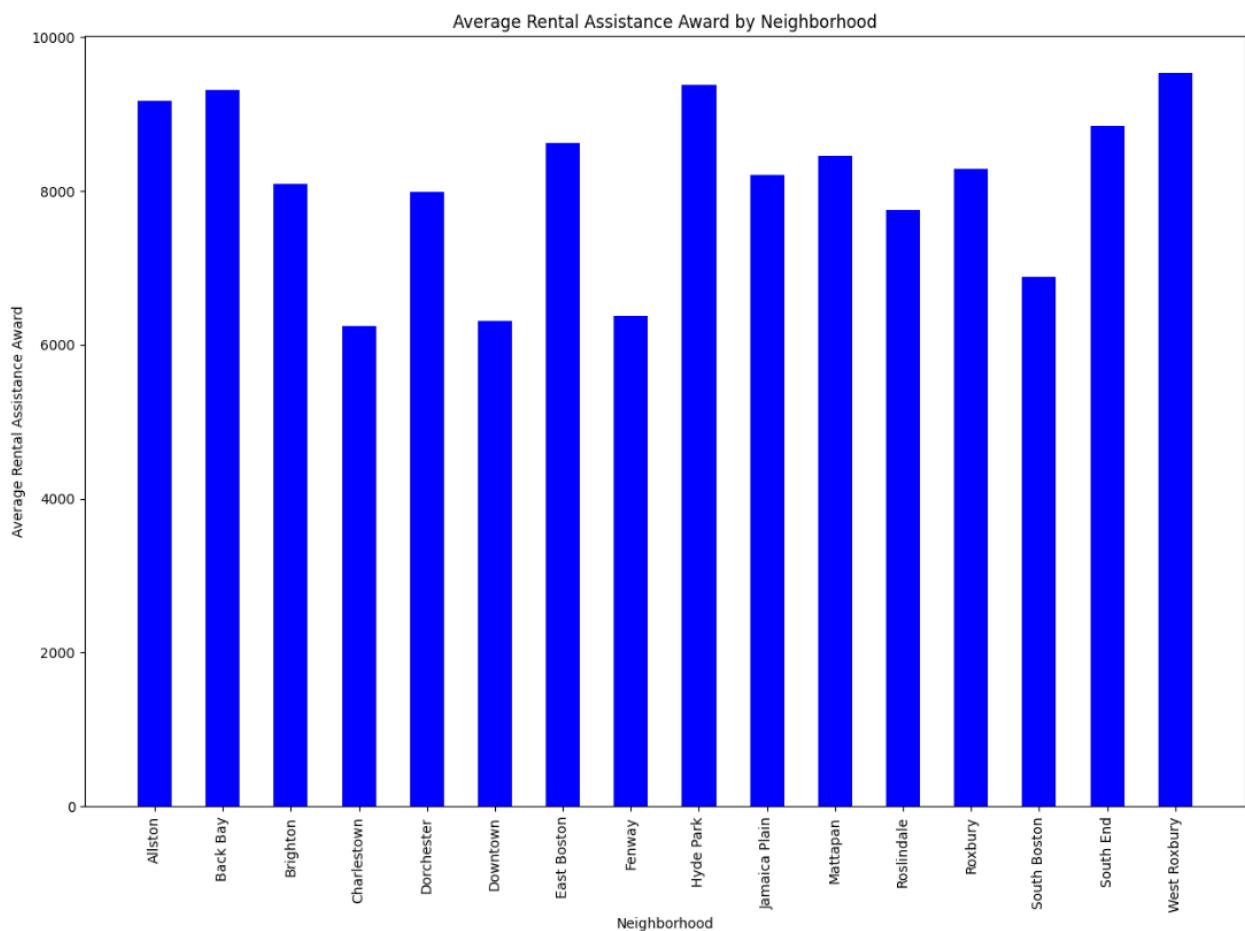
We also analyzed the rental assistance funds, which were separated into a dataset holding data for October 2020 - March 2021 and a second dataset for April 2021 - September 2021.

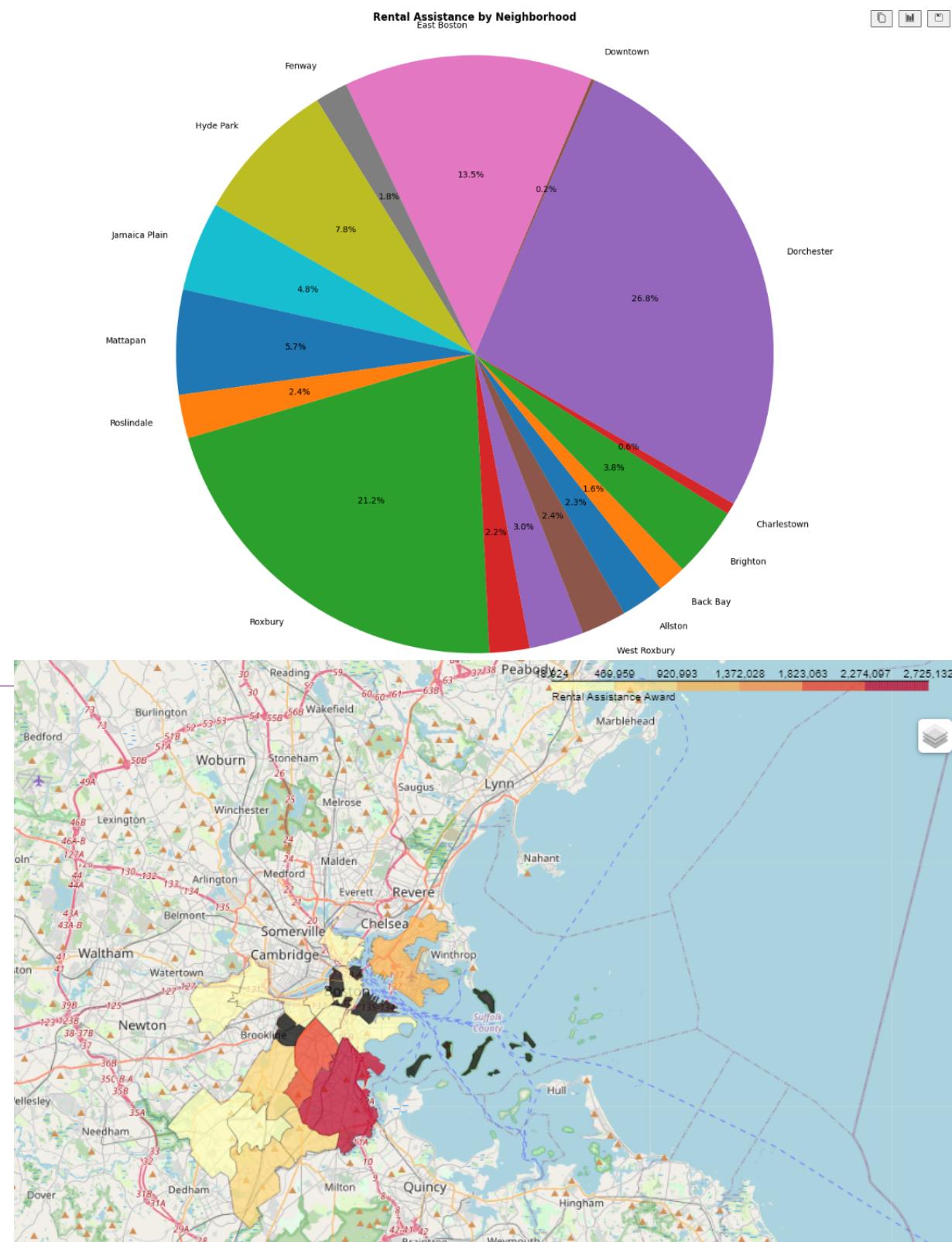
Oct 2020 - March 2021:





April 2021 - September 2021:





From these visualizations, we can conclude that from October 2020-September 2021, the rental assistance funds were equitably distributed to the neighborhoods that we identified as target neighborhoods. For this time period, Dorchester, Roxbury, and East Boston received the largest proportion of rental assistance funds.

Additionally, we can see that the average assistance awarded to a household is approximately equal across all neighborhoods in Boston. There was an average award of \$4561 and standard deviation of \$1364 for October 2020 - March 2021 and an average award of \$8281 and standard deviation of \$4568 for April 2021-September 2021.

Results obtained / questions answered

We wanted to answer the question “Where are the city’s economic development licenses? Which communities are benefitting? Which communities are being left out?” We found that most economic development licenses, including liquor, food, and common victualler, which every establishment that has the capabilities of cooking, preparing, and serving food in-house needs, end up going to Boston and the second most go to Dorchester. This neighborhood benefits a lot because they get more money flowing into their community. With more businesses that have more licenses these cities can attract more people to their city and therefore more money.

We also found that business grants throughout the pandemic have largely been allocated to Dorchester. However, the normalization of the data shows that, while Dorchester receives the highest number of licenses and business grants, they are proportionally not receiving as much as other neighborhoods. In fact, none of the target neighborhoods are receiving an equitable amount of aid from the city. This applies to aid in the form of business licenses, business grants, and capital investments.

Rental assistance, however, has been equitably distributed throughout the pandemic to the neighborhoods that we identified as target neighborhoods.

Interpretation + limitations of results

Throughout the progress of this project the limitations we went through were that the data was difficult to analyze all together since a lot of the data is not really related and there are a lot of empty columns/missing data. The datasets that we analyzed are very different and each serve a

specific purpose for the city of Boston. Many have details associated with their specific purpose but are unrelated to any of the other data. It was difficult to ensure that we were including all the necessary details from each dataset in our analysis.

Furthermore, some datasets categorize their neighborhoods differently. For example, the Harbor Islands is a very significant recipient of Capital Investments, but is not part of the census data for the city of Boston. Furthermore, some of the Capital Investment neighborhoods are grouped together, for example Mission Hill and Roxbury are grouped together in the Business Grants dataset, but Roxbury also has its own independent category. This makes analysis difficult, as Mission Hill is not a target community, so should investments in the “Mission Hill/Roxbury” neighborhood be recognized as equitable investments?

Specifically for the Capital Investments dataset, many of the projects were categorized under “Various Neighborhoods” or “Citywide”. This makes it difficult in our analysis to determine whether the funds are hitting the target communities.

Challenges faced

One challenge that we faced was the limited access to data. Our initial idea for the extension project was not able to reach fruition because we were not able to access data of capital investments from years prior. From the survey data, we can see that many Boston residents also want a greater access to the data of the city’s budget.

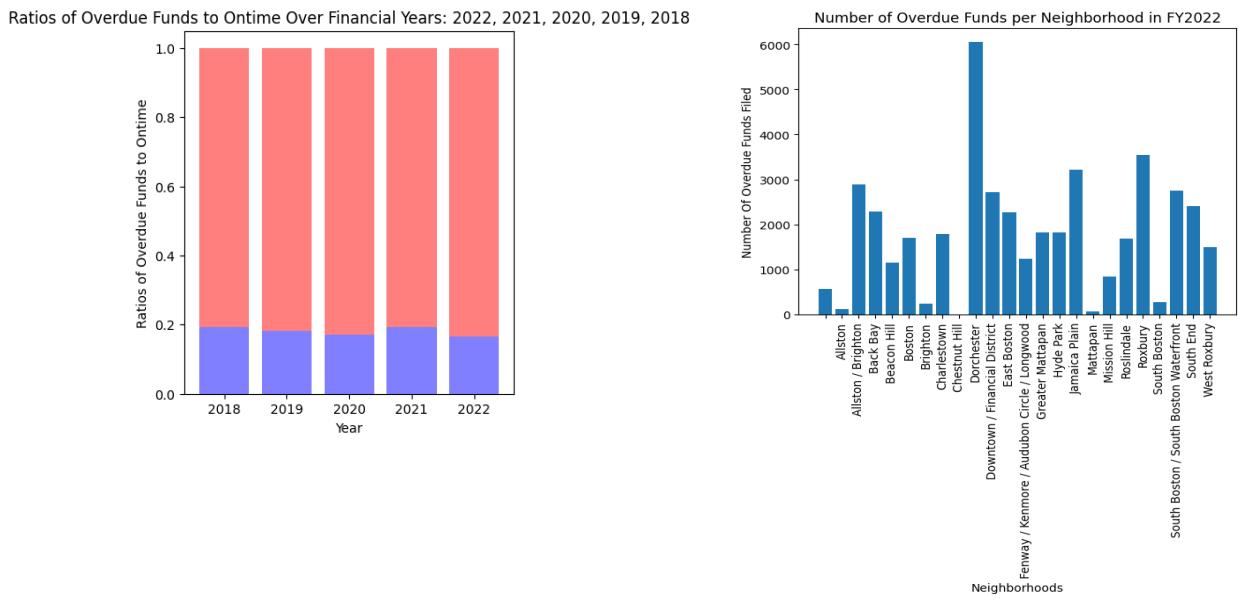
Each dataset had some formatting differences, so we have to accommodate for that. For example some data about costs start with a ‘\$’ whereas others don’t, so when merging the data we had to be careful and watch out for inconsistencies such as that. The rental assistance data was broken down by zipcode rather than by neighborhoods, so to maintain consistency with our other data, we had to determine which zip code aligned with which neighborhoods and populate the dataset based on that.

Additionally, we struggled to normalize the data. As stated above, there was some inconsistency with how each dataset labeled each neighborhood. Neighborhoods were sometimes grouped together and some data included sub-neighborhoods/neighborhoods that were not included in the census data. Because of this, knowing the population of the area that a specific dataset was referencing was difficult. For example, The Capital Investments dataset connected Allston and Brighton into one Allston/Brighton neighborhood. To normalize this data, we added the population of these two neighborhoods together. It also included the Harbor Islands, whose population we collected from: <https://www.city-facts.com/harbor-islands-long-island-boston>.

Explanation of the Extension

Originally, our team planned on analyzing the behavior of the city's expenditure by department. We intended on looking at historical Capital Investment data and determining to what extent a change in leadership impacts the behavior and spending of each department. However, as mentioned under challenges, we were unable to get access to the appropriate data and thus proceeded with the 311 extension project.

On reviewing the data, we noticed that each request had a column, categorizing the request as either 'ONTIME' or 'OVERDUE'. After walking through the data, we decided to calculate the number of overdue requests per neighborhood. Doing so will help us get a better understanding of which neighborhoods may not be getting sufficient funding, since their service request never got fulfilled. We had data available for years 2011 to 2022, but decided to only use data from the past 5 years to ensure results were as accurate and up to date as possible.

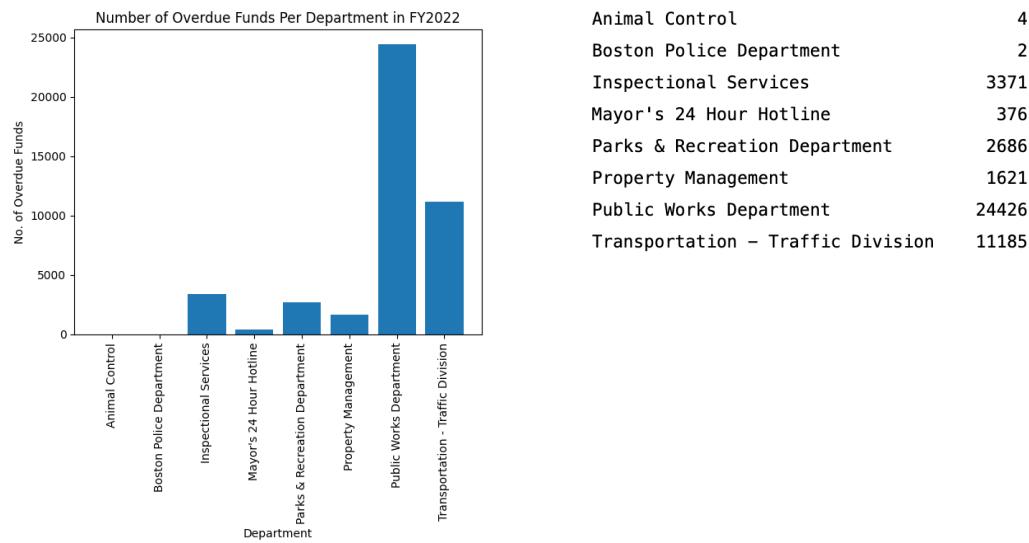


Suggestions for the future of this project/ Solutions

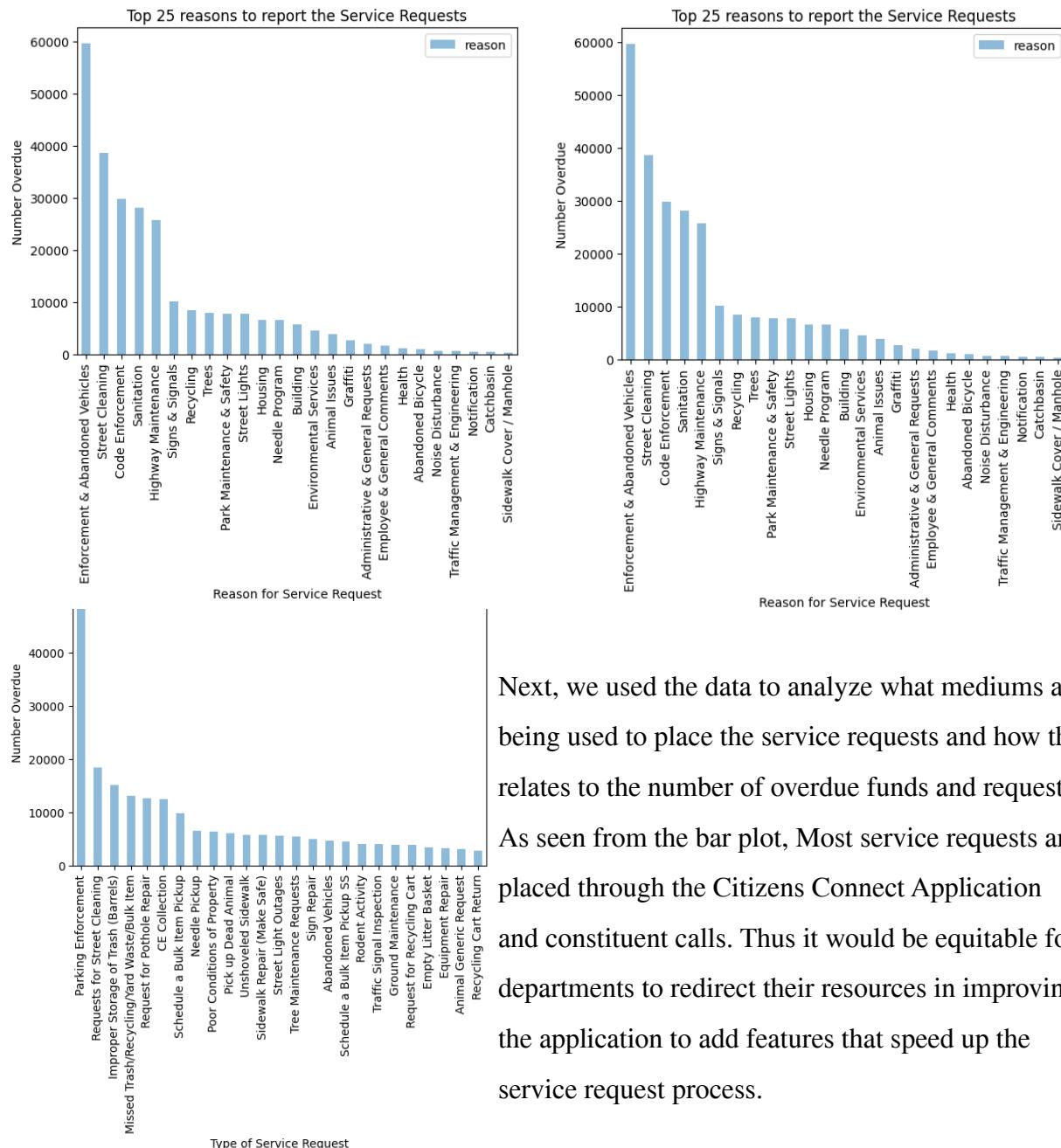
To pose suitable recommendations and solutions for this project, our team decided to incorporate the needs of the Boston residents and analyze the departments that they believe require more funding. To do so, we made use of the 311 Service Requests data.

As seen in the above graphs, we can use the latest 2022 data to analyze which neighborhoods had the maximum number of overdue service requests. As seen through the data, Dorchester seems to be the neighborhood that has the greatest number of overdue service requests. To make suitable predictions for a more equitable fund distribution, more funds can be distributed to areas with higher number of overdue service requests, such as Dorchester and Roxbury.

We also used the data to figure out which departments need most funding by analyzing the proportions of overdue funds to the number of service requests reported to each department. From the bar plot we can see that the public works department has the largest number of requests reported and the largest number of overdue requests. Thus, it would be equitable to provide more funds toward the public works department. We can also see that the Boston Police department and Animal Control department have the least number of overdue service requests. This could mean that the resources provided to these departments are already equitable and do not require major changes.



We noticed that the number of service requests placed are much lower than we actually thought by taking into account the ‘case_status.’ We found that a lot of cases that are reported ‘OVERDUE’ are also marked ‘closed.’ This could be because a lot of the reported cases might be invalid or unfixable. Below we can see the difference in the number of overdue cases after accounting for the cases marked close, which gives us a better understanding of the nature of the requests. For example, before accounting for closed cases, we see that the Public Works department has 24426 overdue cases; but after accounting for closed cases, we see that 11830 of those cases are closed.



Next, we used the data to analyze what mediums are being used to place the service requests and how that relates to the number of overdue funds and requests. As seen from the bar plot, Most service requests are placed through the Citizens Connect Application and constituent calls. Thus it would be equitable for departments to redirect their resources in improving the application to add features that speed up the service request process.

source	
Citizens Connect App	26411
City Worker App	3331
Constituent Call	12737
Employee Generated	734
Self Service	458

Feedback From Client

The City Councilor suggested that since we identified Dorchester as the neighborhood that needs the most money allocated to it, we should take into account its size and diversity. She noted that some parts of Dorchester are wealthier and have a larger white population than other parts, so we should analyze each town differently. This will allow us to get an even more accurate view of which parts of Dorchester need the most attention financially.