City of Boston: Budget Analysis

CS506 Fundamentals of Data Science - Team A
Lucia Kisova (lkisova@bu.edu) - Class of 2024
Toby Ueno (tobyueno@bu.edu) - Class of 2025
Deep Patel (pateld2@bu.edu) - Class of 2025
Thian Amarasekera (tamarase@bu.edu) - Class of 2025

INTRODUCTION

Project Description:
The City of Boston annually collects data on its operating and capital budgets to allocate resources for essential services and infrastructure development. The operating budget encompasses day-to-day expenditures supporting vital roles such as teachers, police officers, and firefighters, along with services like housing, recycling, and transportation. On the other hand, the capital budget, funded through various sources including bonds, city funds, and grants, focuses on acquiring or enhancing physical assets and financing associated projects. The current project is dedicated to a comprehensive analysis of budget distribution across different departments, neighborhoods, and aims to discern trends within funded projects. By scrutinizing datasets covering department and category spending, as well as the proposed operating budget for FY24, the analysis delves into the evolution of funding patterns from 2021 to 2024. The investigation includes a breakdown of spending allocation among diverse departments and programs, an assessment of financial disbursements to various locations, and a calculation of per capita expenditure for different programs.

Project Goal:
This data-driven analysis will provide invaluable insights into the dynamic landscape of budgetary allocations, shedding light on evolving priorities and facilitating informed decision-making for the City of Boston.

Problem Statement:
From the last deliverables, we have been working on analyzing the datasets provided as well as the extension census dataset data from the City of Boston in order to answer all key base questions outlined in the base project as well as our own questions for analysis in the Extension Project. We have been given data regarding the department and category spending for the City of Boston, as well as the proposed operating budget for FY24. We worked on finding out how money has been spent for the years 2021, 2022, 2023, 2024, by breaking down how spending has been allocated to different departments and programs, and how much different locations receive. Additionally, we are looking at how much money is spent per capita for all the different programs. Lastly, we are analyzing a subset of 2020 census data from the City of Boston website containing population statistics for neighborhoods including breakdowns by race/ethnicity, age, household size, university housing, and some other miscellaneous information. We aimed to investigate neighborhood spending per capita to inform equitable urban planning and development strategies for local policymakers and businesses.

Data Collection/Cleaning:
Our data was pre-collected from the City of Boston website for the Operating Budget Data and Planned Capital Budget Data; no extra steps for collection were necessary beyond downloading the associated CSV files. For the extension project we have the subset of the 2020 census data from the City of Boston website for the data collection. Each column was identified via a code, such as P0020001, which was manually cross-referenced with a Census summary file to extract the meaning of the code. Cleaning and preprocessing this dataset involved renaming the columns to carry semantic value instead of these arbitrary codes, as well as typical data type conversion and removing or rearranging out-of-place rows and columns (such as that of the Neighborhood, which was moved to the index of the dataframe). One limitation to note is that the census data is usually collected every 10 years, as a result the 2020 subset (as previously mentioned) is the most recent sample. This creates a hindrance for the insights and results obtained during the analysis steps.

To clean the data, the main step was to get rid of missing values, particularly in numeric fields. Numeric fields had the string "#MISSING" instead of a NaN value, sometimes with whitespace surrounding the string, so each numeric column was first treated as a string, stripped of whitespace, and then all rows with any values equal to this "missing" string were dropped. Columns were then safely cast to numeric values.

Project Impact:
By utilizing data to create visual representations of how the Boston city budget has been allocated across various different departments, cabinets, categories, etc, it can become easier to identify trends among budget allocations. By using these visuals to identify trends, areas where the budget can be reduced can be more easily identified, and areas where more budget needs to be allocated can be more easily found. In addition, it can become easier to identify areas where there may be a bias in the allocation of funds to certain departments or neighborhoods.

BASE PROJECT

Note: An important note is that the Execution of Courts department and Annual Audit costs reported a significant percent increase in funding.

Hypothesis: We hypothesize that these two are correlated. We hypothesize that the city is increasing the funding in these departments to potentially increase revenue in future years.

Key Question #1: **How has spending per department changed over time?**

To address the first base question, we look at the average spending across all the departments. We found that on average, the city budget increased by **5.59%** per year. Furthermore, we wanted

to find the percent increase in the budget spending across the years. We noted that the budget had a **6.56%** increase from 2021-2022, a **3.56%** increase from 2022-2023, and the budget is projected to increase by **6.65**% from 2023-2024.

Next, to further analyze the spending across departments over time, we found which departments received the largest percentage increase in terms of budget spending per year over the four year period. The three departments which gained the largest percent-increase in funding per year over this four-year period were:

- Execution of Courts (under the Finance Cabinet), with a **136.81**% increase
- Annual Audit Costs (under "Other"), with a **133.88**% increase
- Reserve for Collective Bargaining City (also under "Other"), with a **117.21**% increase

An important note is that this excludes the OPAT cabinet, which received no funding in 2021 since it was conceived in 2022.

Conversely, to better understand how spending differed across departments over time, we wanted to identify the largest percentage decrease in terms of budget spending per year over the four year period. The three departments which lost the largest percent-decrease in funding per year over this four-year period were:

- Office of Finance (under the Finance Cabinet), with a **20.12**% decrease
- Suffolk County Sheriff Dpt. (under "Other"), with an **8.32**% decrease
- Library Department (under the Human Services Cabinet), with a **4.89**% decrease

It is important to draw attention to the fact that along with these three departments, The Police Department and the Central Fleet Management were the only other two departments to report an average yearly decrease in funding over this period. The remaining 61 other departments reported an increase in budget spending per year over the four year period.

To further examine the first base question for the project, we conducted analysis on the data to produce visualizations in order to draw meaningful conclusions from the data.

Figure 1: This line chart provides insight into the data showing the increase in funding for the departments. Based on this visual, we can see there is a general increase year over year by looking at the percentage increase from each year. It is important to note that while there evidently are some outliers, however, most departments show an increase in spending.

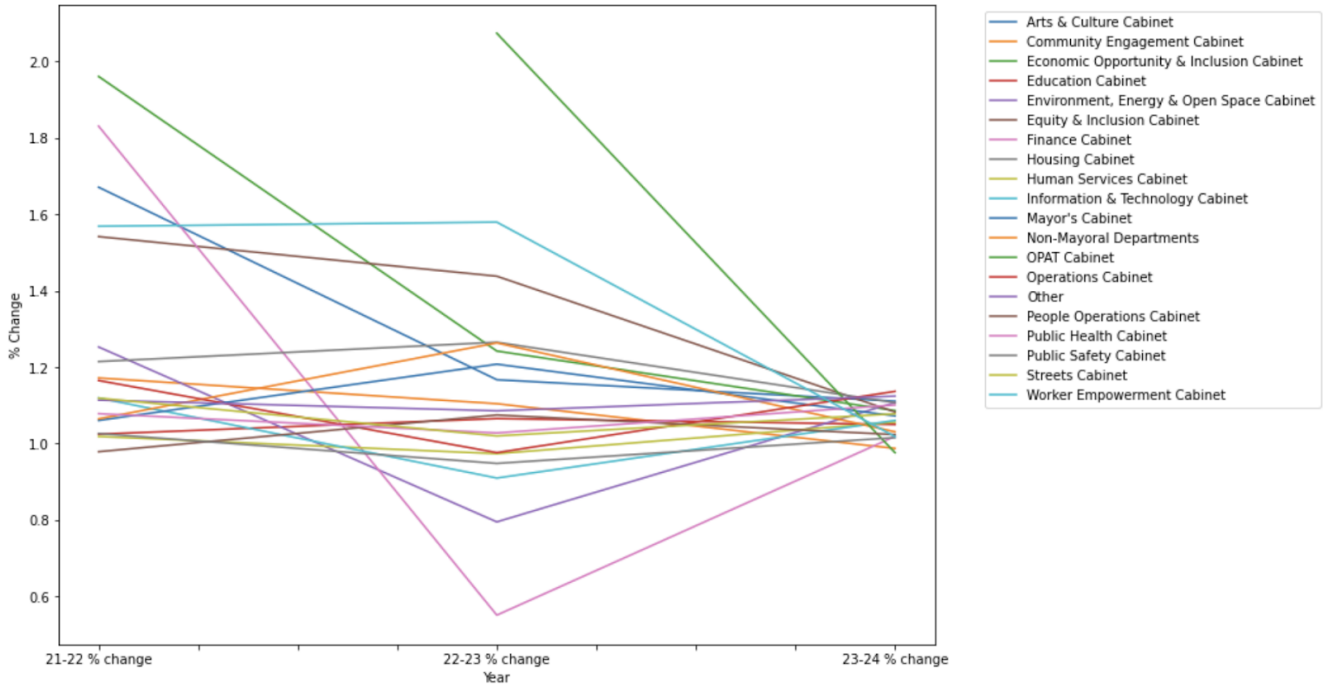


Figure 2: This bar chart shows the expenditure of the various cabinet departments in the city of Boston for the 2023 fiscal year (actual + project).

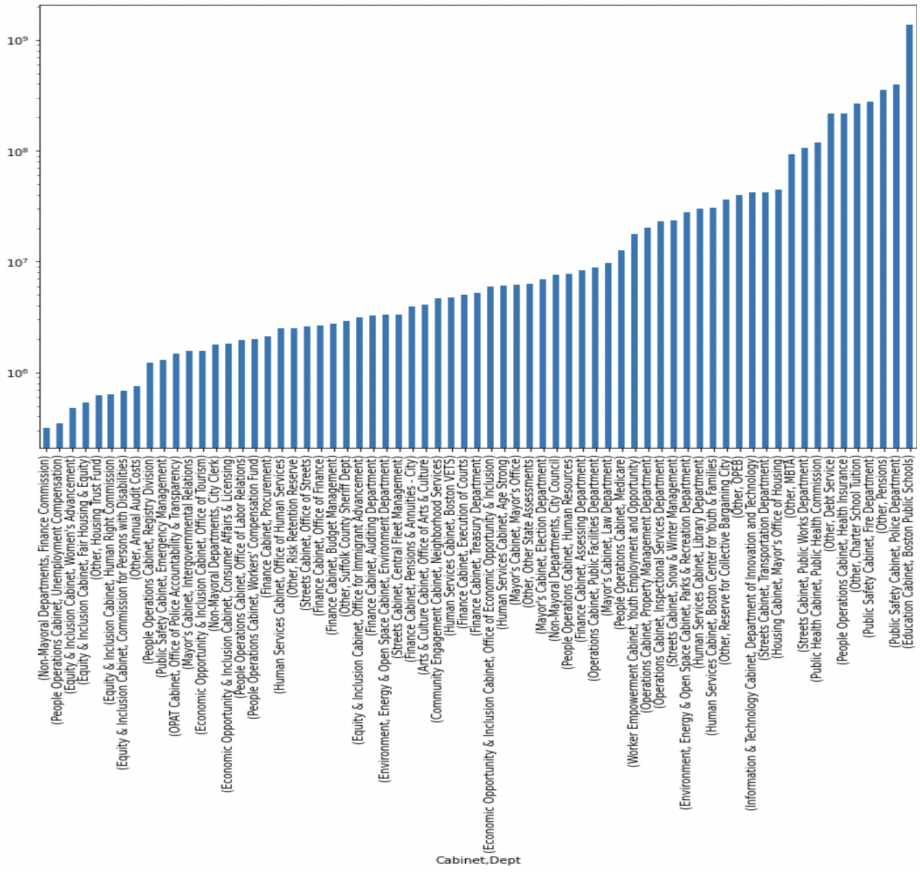
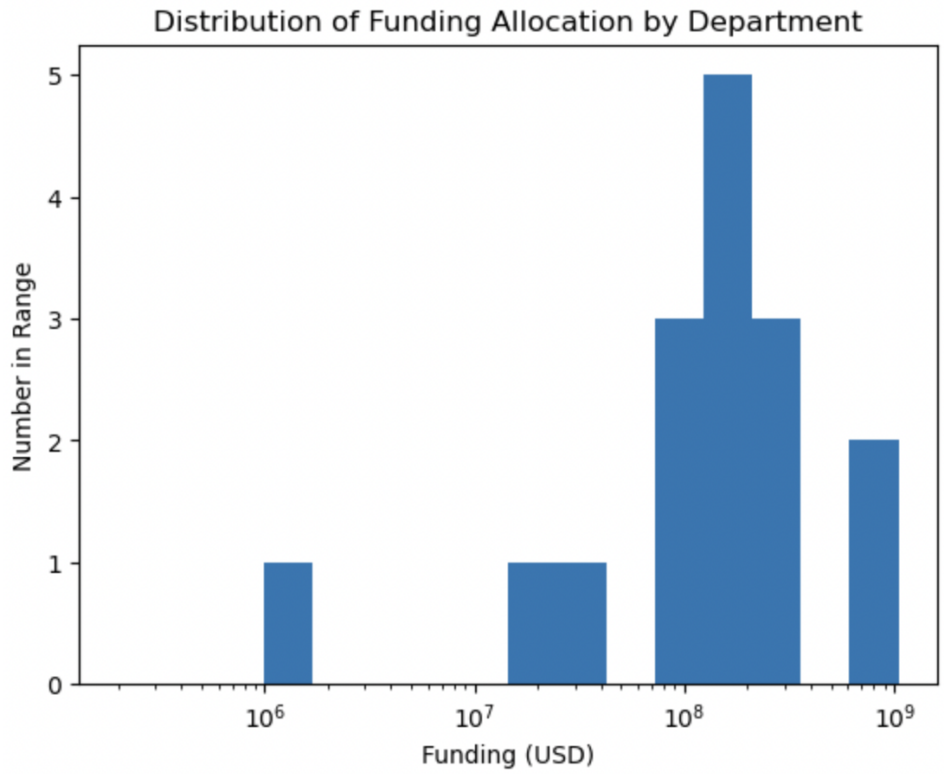


Figure 3: To give further insight into the funding by departments, we created a similar bar chart to show the projected distribution of funding for the year 2024. The distribution displays the amount of funding that each department received. As seen below, there are many departments that are receiving a large amount of funding. On the other hand, there is only one department as an outlier that is predicted to receive relatively low funding. The rest of the departments are expected to receive funding between $10 million and $1 billion. It is noteworthy to draw attention to how most of the departments will receive around $100 million in funding.



Key Question #2: **How has spending per program changed over time?**
To address the second base question, we looked at the department budgets that had the largest percentage increase and decrease in funding over the four-year time period of the dataset. We found that this was sufficient information to properly answer the second base question.

Initially, we wanted to identify the programs with the highest percentage increase over the period of time. The three programs which gained the largest percent-increase in funding over this four-year period were:

- Trust (under the Treasury Department), with a **4854.92%** increase
- Small & Local Business (under the Office of Economic Opportunity & Inclusion), with a **656.51%** increase
- Reserve for Collective Bargaining City (under department of the same name), with a **444.38%** increase

Some important things to note would be that The Trust project can be viewed as somewhat of an outlier, since its budget in 2021 was relatively low ($631.62). Similar to the analysis of the first base question, the analysis excludes the OPAT cabinet, which received no funding in 2021 since it was conceived in 2022.

Now, we wanted to identify the programs with the highest percentage decrease over the period of time. The three programs which suffered the largest percent-decrease in funding over this four-year period were:

- Affirmative Action (under Human Resources), with a **97.31**% decrease
- Consumer Affairs (under Consumer Affairs & Licensing), with an **56.63**% decrease
- Alterations & Repair (under the Property Management Department) with a **54.68**% decrease

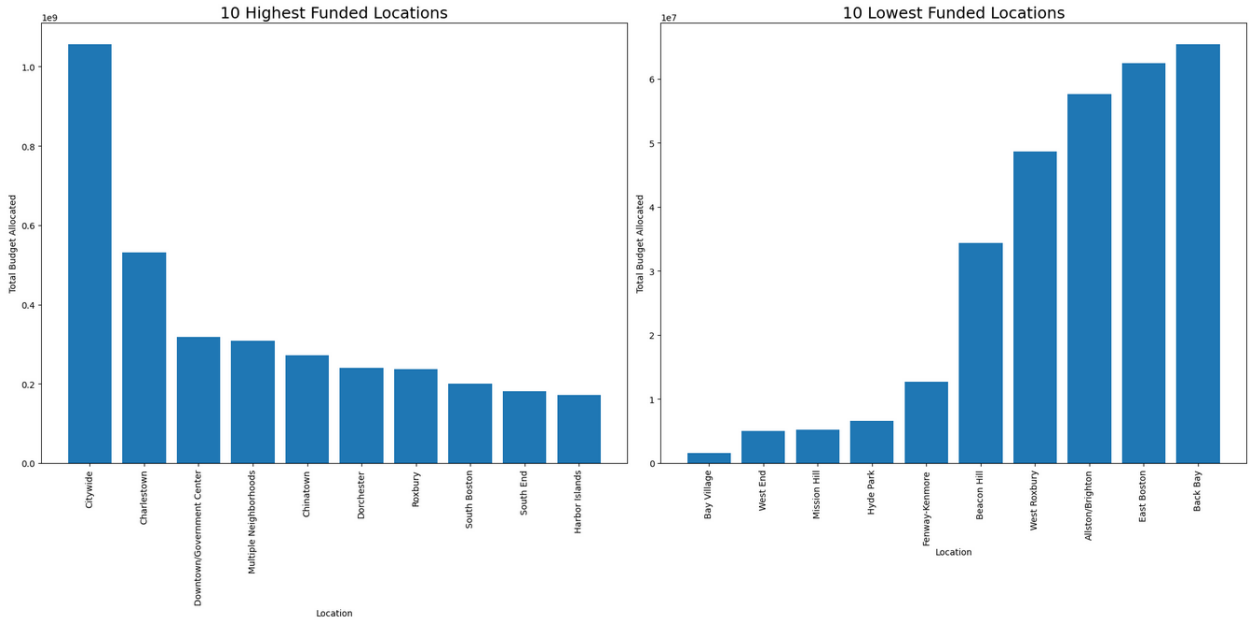
Based on the information on the three programs which experienced the largest percentage increase and decrease in funding over the four-year period, we also wanted to look at the overall funding for all 184 projects. We investigated that 151 projects received an increase in funding, which is **82.07**% of the total number of projects. Conversely, 33 projects received a decrease in funding over the time period which is **17.93**% of the total number of projects.

Key Question 3: **How is spending allocated to different locations in the city?**

To address the third base question, we looked at spending distribution with regards to different locations in the city. We primarily focused on finding the funding in each neighborhood and this got us curious to find more information about the spending per capita in neighborhoods in Boston, which will be investigated thoroughly in the Extension Project. For now, we are focusing on the allocation of spending in different parts of Boston. In order to tackle this question, we found the neighborhoods that are most and least funded.

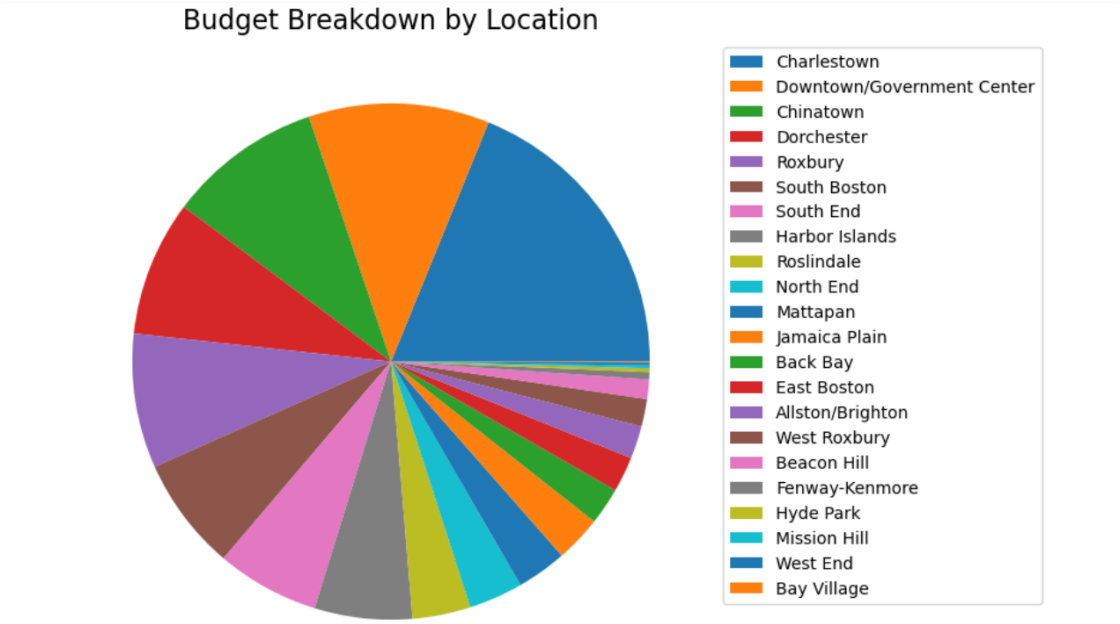
Excluding citywide projects, Charlestown was by far the most funded neighborhood, with a total budget of $**531.2** million. On the other hand, Bay Village was the lowest funded neighborhood with a total budget of only $**1.53** million. The difference between the funding for these two neighborhoods is incredibly significant. To further examine the allocation of spending in different parts of Boston, we looked at the percentage breakdown of the most and least funded neighborhoods in relation to the total funding for neighborhood-specific projects. We found that Charlestown's budget is **18.84**% of the total neighborhood-specific project budget, whereas on the other hand, Bay Village's budget is only **0.05**% of the total neighborhood-specific budget. This further emphasizes the drastic difference between funding allocation in different neighborhoods in Boston.

Figures 4 & 5: These two bar charts show the allocation of funding among the top 10 highest funded and 10 lowest funded neighborhoods in the city.



There were a total of 22 neighborhoods that had projects that were included in the Capital Budget Plan. In the extension project, we broke down the demographics of different neighborhoods and analyzed trends in higher- and lower-funded areas.

Figure 6: This pie chart shows a full breakdown of funding allocation to each of the 22 specific neighborhoods. However, this excludes citywide and multiple neighborhood projects.



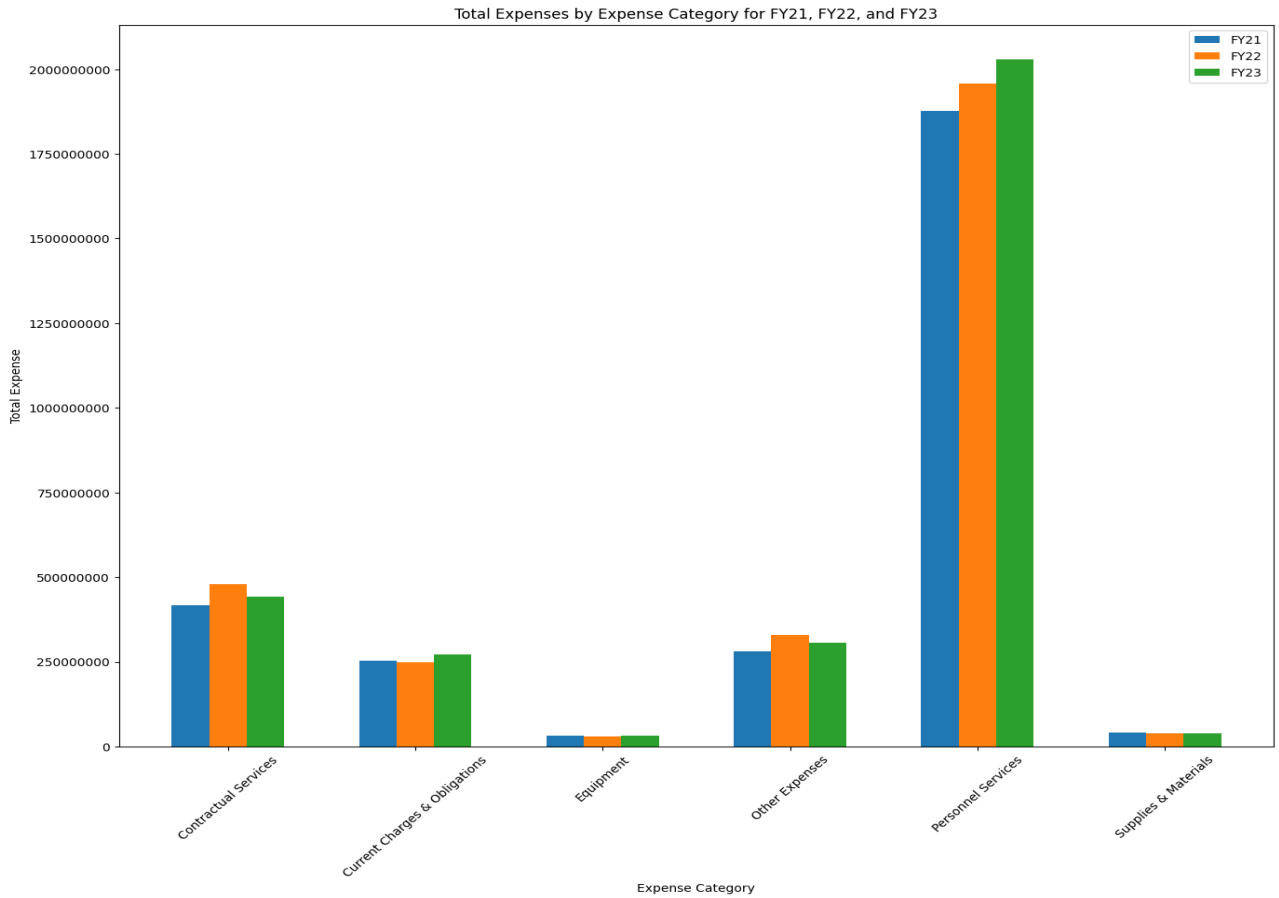
Key question #4: **How is spending broken down by budget category?**

To address the fourth base question, we wanted to find the main expense categories and, from that, determine how spending has been broken down by budget categories. We wanted to primarily focus on finding the largest and lowest expense budget categories over the three year-period.

There are 7 main expense categories that the City of Boston has. These are: Personnel Services, Contractual Services, 'Supplies & Materials, Current Charges & Obligations, 'Equipment, Other Expenses, and Fixed Expenses. From these, the lowest spending category are the Fixed Expenses.

For the fiscal years 2021, 2022, and 2023 these expenses on average have been about $**191,092,897**. The Personnel Expenses are by far the largest expense category each year. Personnel expenses have been consistently above $**1.8** billion dollars per year for the past 3 years. One interesting thing to note is that it has also been the one of the only expense categories where the expense has risen for each of the past 3 years. All other categories, besides Fixed Expenses, have had some degree of fluctuation on a year to year basis. This can be attributed to a couple economic factors such as an expanding workforce and the resulting increase in wages to compensate for inflation and high cost of living adjustments.

Figure 7: This bar graph depicts the total expense by each of the categories for the past years - 2021, 2022, 2023. It is important to mention that this graph excludes the Fixed Expenses category which is far lower than the rest of the expenses that are available to us in the dataset. As seen, Personnel Services cost the City of Boston upwards of $**1,800,000,000** per year. Whereas, Contractual Services, which is the second most costly expense, took around $**450,000,000** per year. Equipment and Supplies & Materials both cost much less than **$250,000,000** per year, displaying the significance of the disparity between the most and least costly expense budget categories.



EXTENSION PROJECT

Extension Project Goal:
The extension project aims to uncover and analyze trends in neighborhood spending per capita, shedding light on the variations and disparities that exist across different neighborhoods in Boston. The hope from this extension is by delving deeper into the exploration of the third base question, we hope to provide valuable insights that can inform local policy-makers, business strategies, and community initiatives to foster a more equitable and informed approach to urban planning and development regarding the City Budget.

Relevance:
This extension project is important as it can unravel the intricate relationship between neighborhood funding and population density. This could offer a better understanding of how expenditures vary across diverse communities in Boston. The significance lies in the visibility of these expenditures to local residents, making the findings essential for shaping policies and initiatives that are designed to address the needs and preferences of each respective neighborhood. By bridging the gap between spending patterns and population density, the project aims to empower communities, local authorities, and businesses with actionable insights. This

can be important when looking to the future in order to create a more responsive and community-centric approach to urban development in Boston.

Our analysis will delve into pivotal questions surrounding neighborhoods with higher spending per capita, aiming to hone in on trends and similarities among those areas. By investigating whether proximity and smaller population density contribute to heightened spending patterns, the project seeks to validate a preliminary hypothesis. This will provide valuable insights into the spatial and demographic dynamics that shape neighborhood expenditure. These questions serve as the foundation for uncovering patterns that can inform targeted interventions and policies, fostering a more comprehensive understanding of the economic dynamics at the community level in Boston.

Data:
By using the subset of the 2020 census data from the City of Boston website, which provides information on population density with corresponding neighborhood names and population figures we leveraged this comprehensive dataset to conduct analysis of funding across various neighborhoods with different spending per capita. This allowed us to draw correlations between population density and expenditure patterns.

Our data visualization centers around the creation of informative visually compelling tools detailing the spending per capita for each neighborhood and the population density. By employing geospatial visualization, we can provide a clear and accessible representation of spending patterns across the city. This visual tool not only enhances the communication of our findings but also facilitates a more intuitive understanding of the economic landscape in Boston neighborhoods which can aid policy-makers in making informed decisions related to neighborhood development and resource allocation.

Hypothesis: Neighborhoods that contain a larger percentage of older individuals and white individuals have a larger portion of the city budget allocated to them as measured by budget per capita.

Extension Question #1: **What is the population density of the neighborhoods in Boston?**

To address this first question, we looked at the populations of each of the neighborhoods in Boston as well as the average household size by neighborhood. This took an important step in finding the distribution of populations amongst the neighborhoods in Boston, and as a result we can determine the spending per capita. Furthermore, the range of the average household size can give further insight into the population density of each neighborhood.

Figure 8: The bar graph showcases the overall population by neighborhood. As we can see below, Dorchester notably had the highest population by a large margin as it was roughly double that of second-place Roxbury. The third most populated neighborhood, Brighton, is roughly similar to Roxbury. On the other hand, Longwood had the smallest population from all the neighborhoods in Boston. Finally, it is important to note that South Boston Waterfront had a similar population size to Longwood and was the second smallest neighborhood.

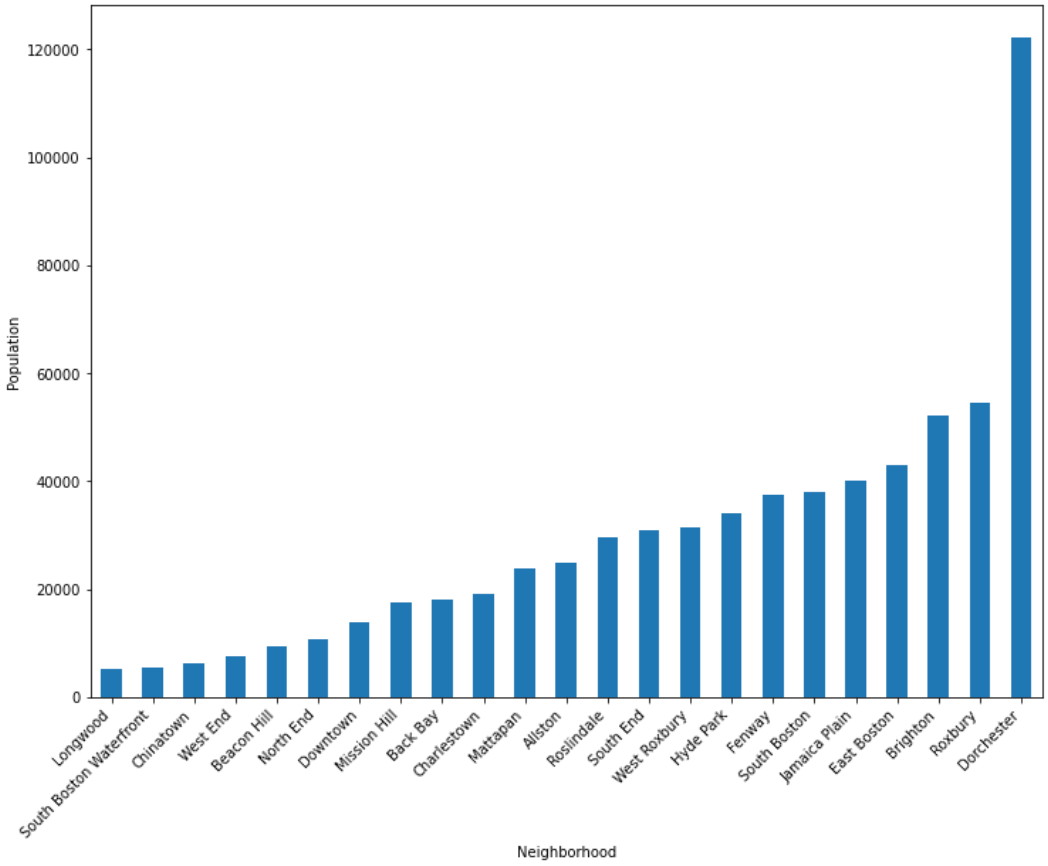
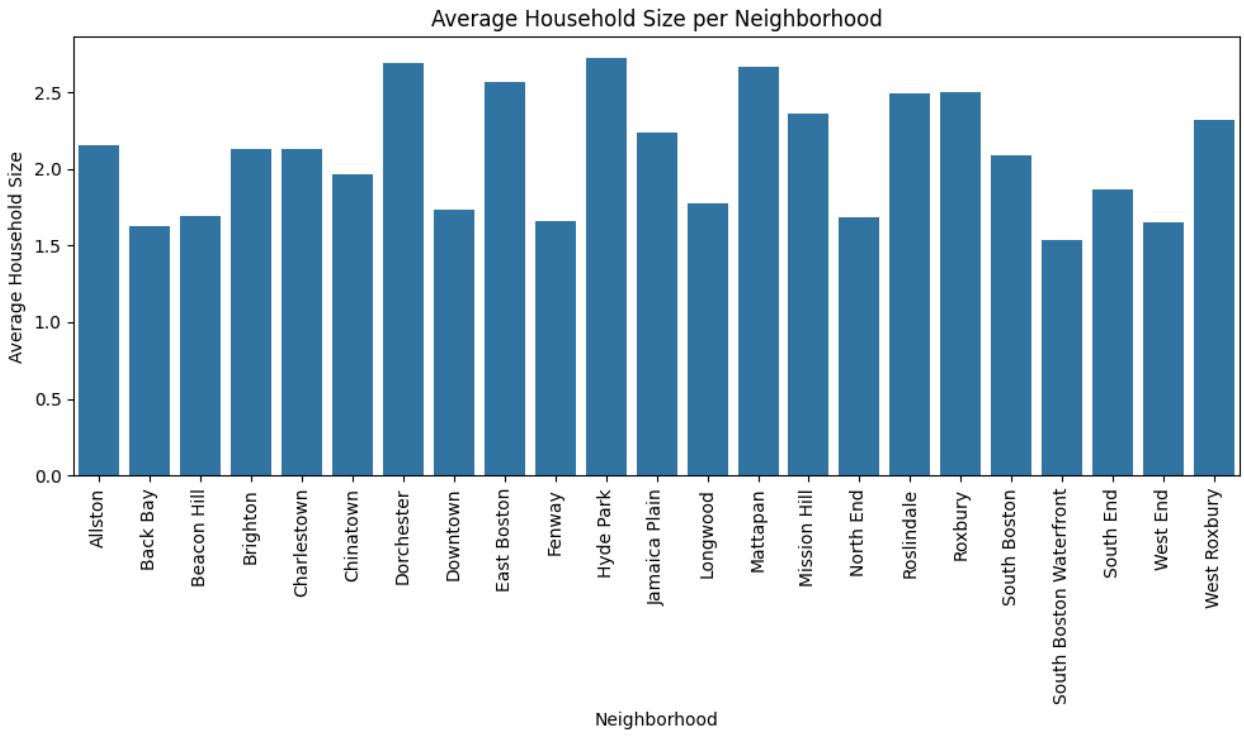


Figure 9: This bar graph displays the average household size by neighborhood. It is evident from the bar graph there does not seem to be any particular neighborhood that is an outlier for this metric. The smallest average household size is 1.53 for South Boston Waterfront, which follows Figure 8 this neighborhood has the smallest population. On the other hand, the largest average household size is 2.72 for Hyde Park. This is interesting to discover as Hyde Park was the 8th most populated neighborhood but has the highest average household size. Therefore, we can infer that Hyde Park has densely populated households. From the bar graph below, we can conclude that the typical range of household size ranges from 1.53 (South Boston Waterfront) to 2.72 (Hyde Park).



Extension Question #2: What are the demographic characteristics of the neighborhoods?

To address the second extension question, we wanted to look at the breakdown of age, race and college students in each neighborhood. This was necessary to consider as we wanted to detail the disparities between neighborhoods when looking at spending per capita. Initially, we saw the breakdown of age groups for every neighborhood. Furthermore, we investigated the racial demographics of the populations. Lastly, since Boston is home to multiple universities, we wanted to see the number of college students in all the neighborhoods that are home to college students.

Figure 10: The following bar graph displayed the population per age group for each neighborhood. The two age groups we created for each neighborhood were 0-17 years old and 18+. In every neighborhood, the number of individuals who are older than 18 is far greater than the number of individuals who are 0-17 years old, which is to be expected. One thing to note is that most neighborhoods have a disproportionately larger number of individuals who are 18+ compared to the very small percentage of the population who are 0-17 years of age.

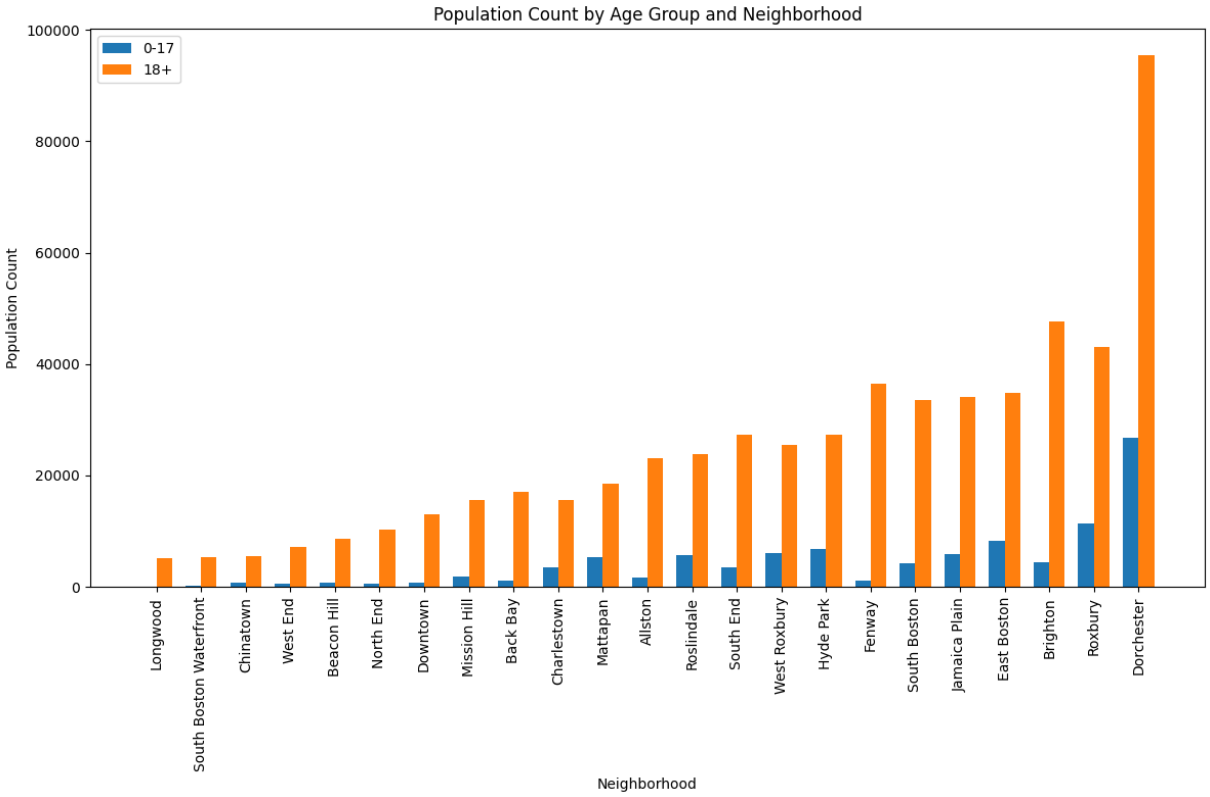


Figure 11: This bar chart depicts the number of college students in each neighborhood of Boston, excluding those that do not have college students. Fenway appears to be a very popular location for students to live, perhaps due to factors such as being located near a large university.

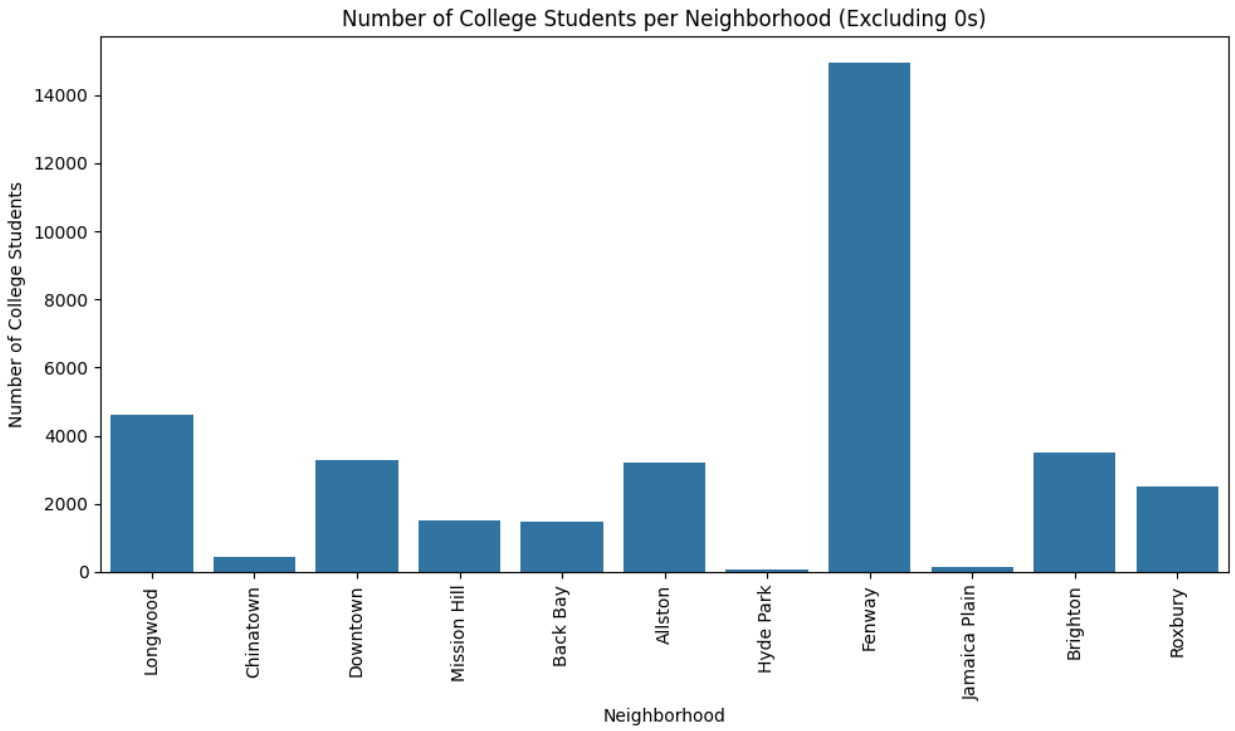


Figure 12: Bar chart showing the race demographics of Dorchester, the neighborhood with the aforementioned largest population. Black or African American residents comprise the largest single racial group of Dorchester residents, at around 35% of the total population. White and Hispanic/Latino residents are about even at around 21% each, followed by Asian/Pacific Islanders and Other (which includes two or more races) at around 11% each.

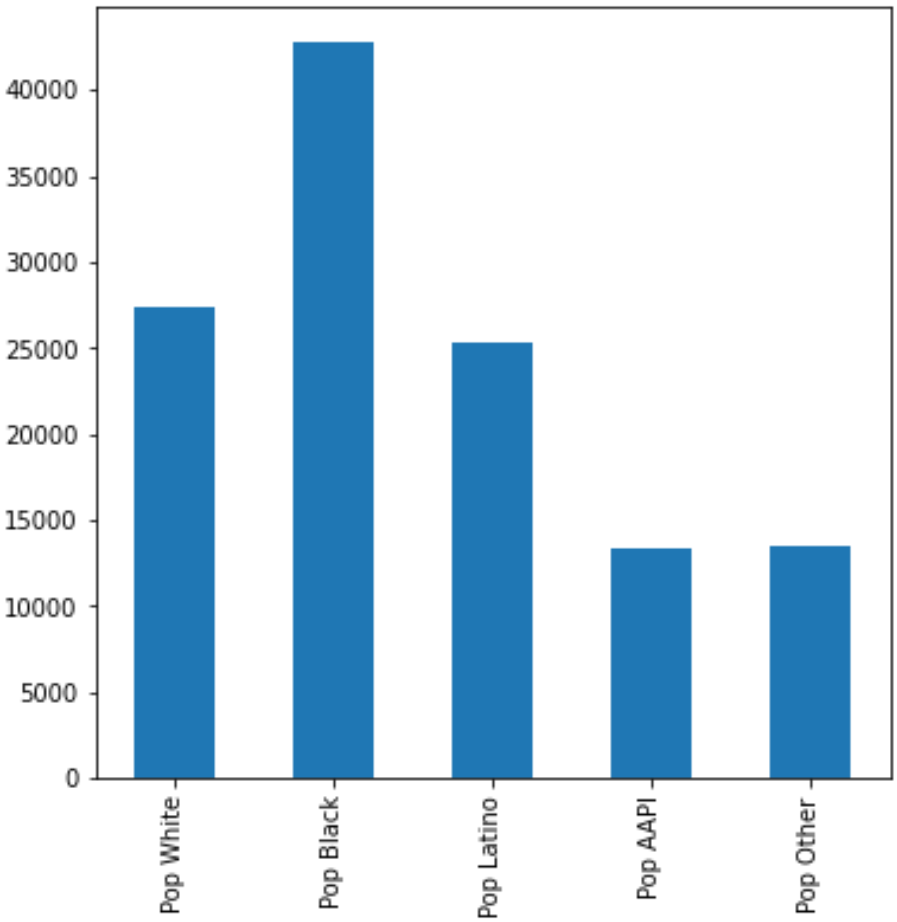
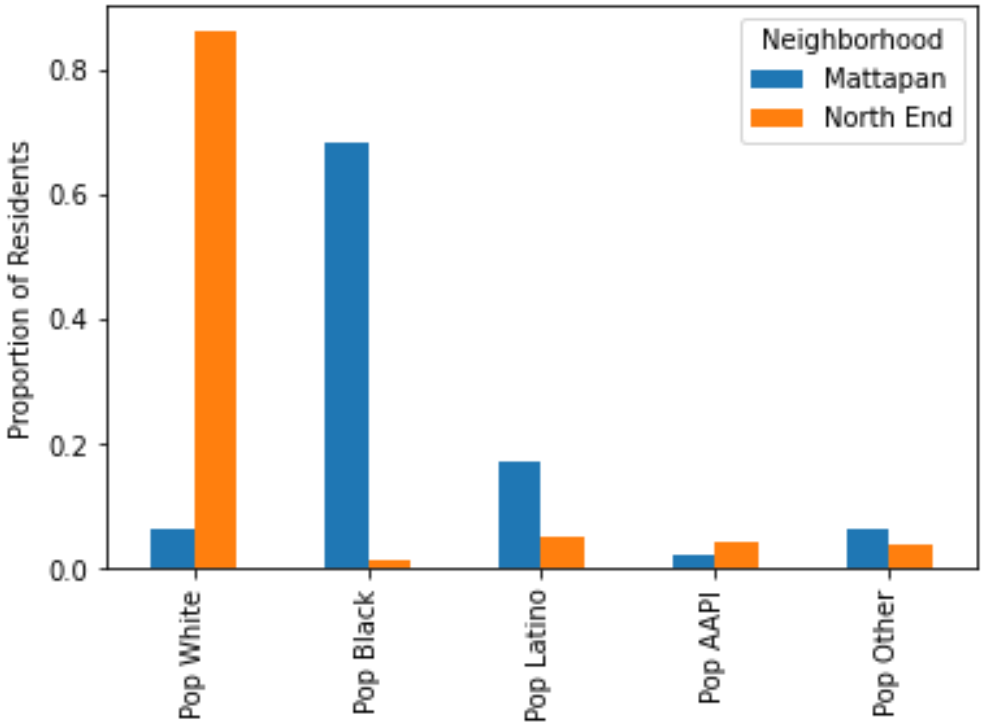


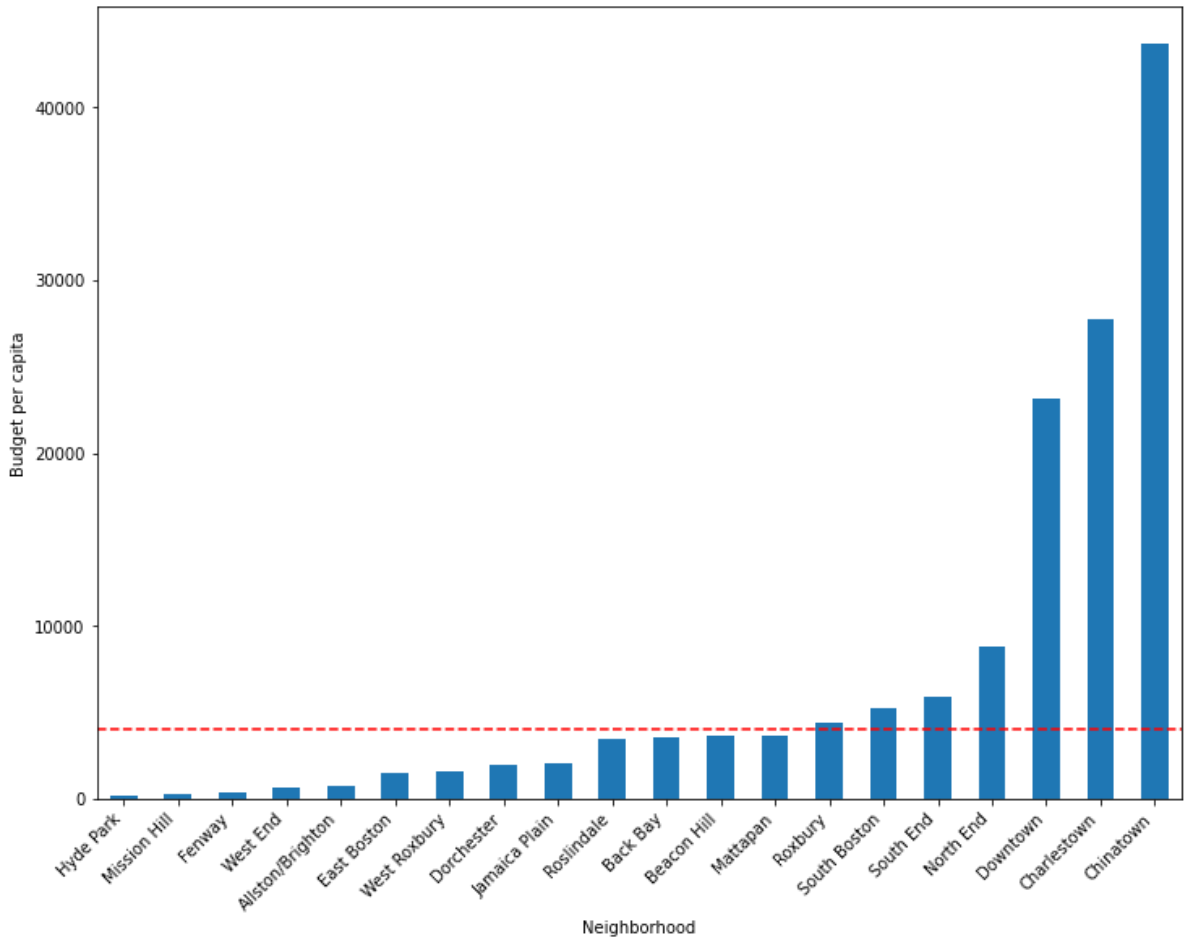
Figure 13: Comparative bar chart between the demographics of Mattapan and the North End, the neighborhoods with the largest and smallest proportions of Black residents, respectively. The y-axis is normalized to proportion of population instead of absolute numbers to account for the neighborhoods being different sizes. Here, we observe an anecdotal inverse relationship between Black and white populations; when one is high, the other is not. Furthermore, the percentage of residents of other races is relatively low for races other than the majority.



Extension Question #3: What is the spending per capita by neighborhood?

To address this question, we found the mean budget-per-capita across all neighborhoods. We investigated the budget spending for each of the neighborhoods as a stepping stone to tackle our overall extension project.

Figure 14: This bar chart shows the neighborhood's project budget allotment per-capita. We can see that Charlestown, Chinatown and Downtown were the neighborhoods that received the most spending per-capita. These three neighborhoods can be considered as outliers and help increase the mean budget spending per capita for all neighborhoods in Boston. The neighborhoods that received the least amount of spending per-capita were Fenway, Mission and Hyde Park. This is interesting to consider because Fenway and Mission Hill are home to many college students. Furthermore, the mean budget spending per-capita across all neighborhoods is outlined in the graph through the red dotted-line at ~$3,980.



Extension Question #4: How does budget allocation differ by neighborhood race demographics?

To address this question, we calculated the budget allocated to each of the neighborhoods in Boston by using the data provided by the Spark! team. In addition, we used the data from the 2020 census of Boston to collect information regarding the demographics of each of these neighborhoods. By comparing these two datasets, we were able to determine whether or not there was a correlation between the race demographics of the neighborhood and the budget per capital funding that the neighborhood would receive. One key thing to note is that from Figure 15, there does not appear to be much of a correlation in funding per capita depending on the fraction of the population that is white in a neighborhood. It can be seen that when the percentage of the white population starts to become greater than about 65%, the budget per-capita for that neighborhood does seem to increase slightly, though this is not a strong trend. Figure 16, however, shows that neighborhoods with larger black populations are among the lowest funded. The same is seen in Figure 17 with latino populations, however there is only one neighborhood with a latino population larger than 30%, so this trend is not as pronounced. There does appear to be a trend that neighborhoods generally do not tend to have large black or latino populations, though the same does not hold for the white populations.

Figure 15: Scatter plot depicting the budget per capita of Boston neighborhoods as a function of the percentage of white individuals in the neighborhood. (0.1 = 10% white, Budget_PC in USD)

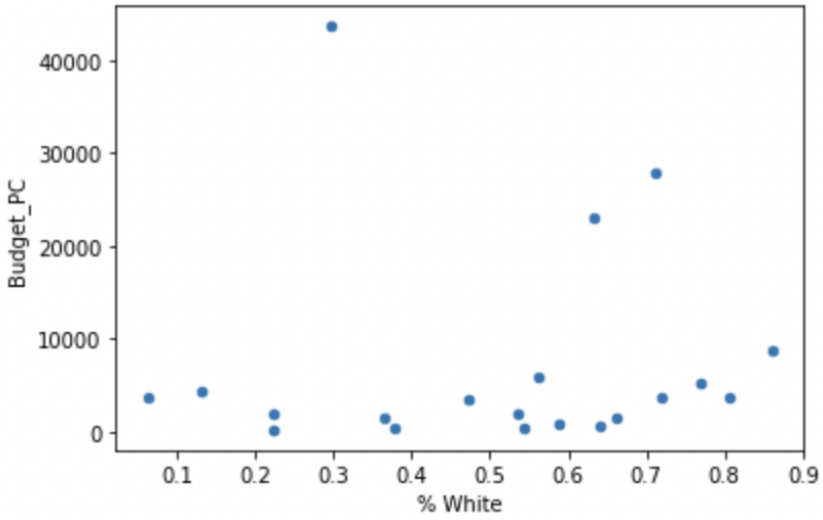


Figure 16: Scatter plot depicting the budget per capita of Boston neighborhoods as a function of the percentage of African American individuals in the neighborhood. (0.1 = 10% black, Budget_PC in USD)

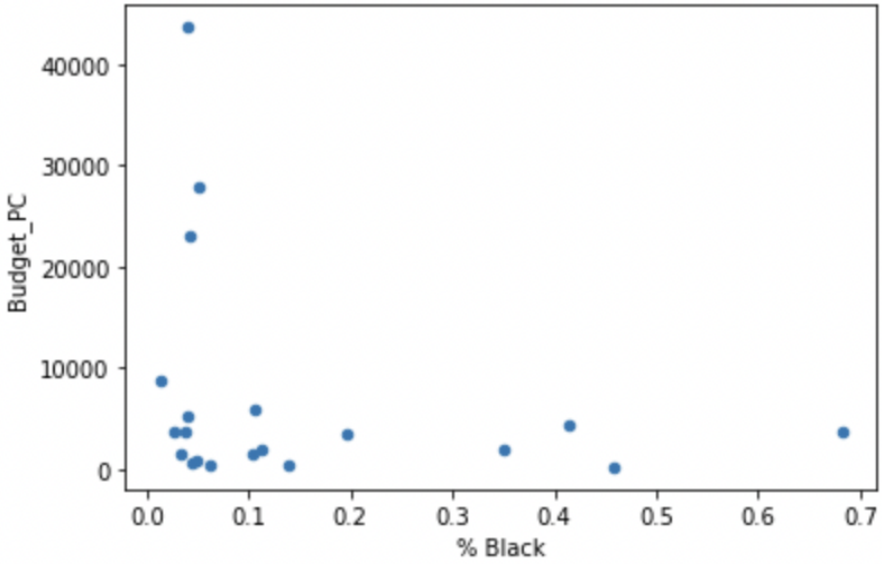
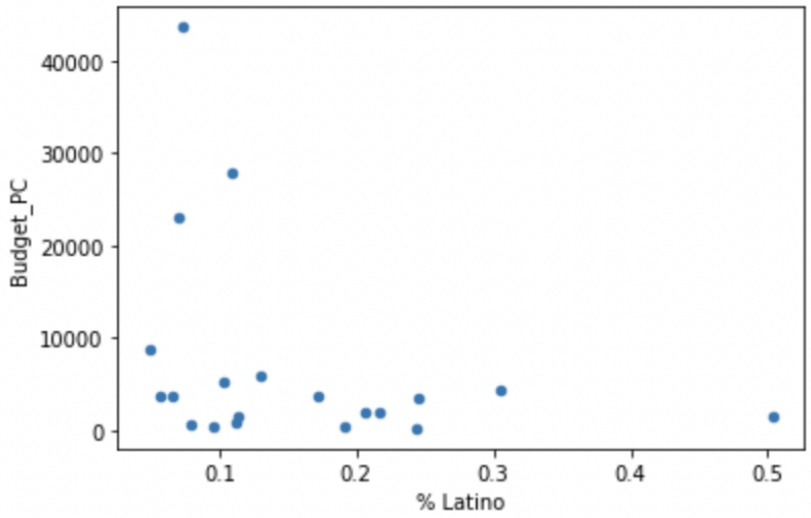


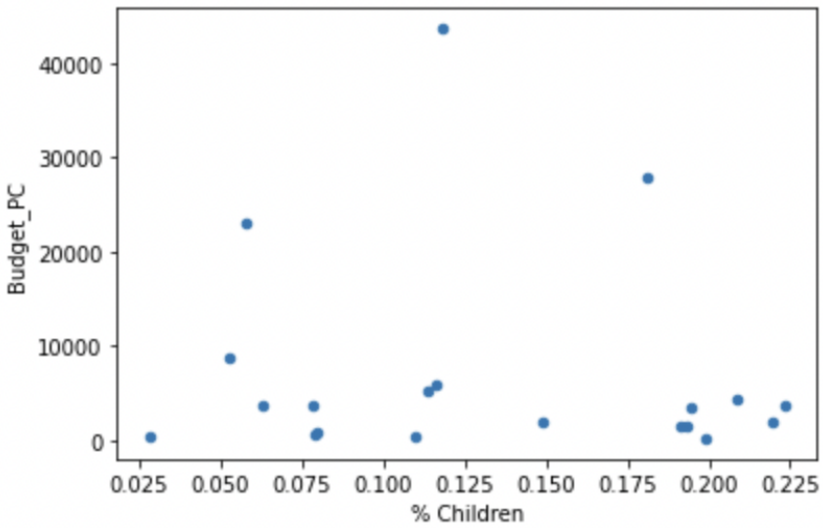
Figure 17: Scatter plot depicting the budget per capita of Boston neighborhoods as a function of the percentage of Latino individuals in the neighborhood. (0.1 = 10% latino, Budget_PC in USD)



Extension Question #5: How does budget allocation differ by neighborhood age demographics?

Initially, we believed that there would be a correlation between neighborhoods with a larger population of children (0-17 years) and the budget allocation to that neighborhood. From figure 18, it is evident that no such relation exists. There do exist some outliers, however the vast majority of neighborhoods have a relatively similar budget allocation per capita and the population of children does not seem to play a role in the budget allocation.

Figure 18: Scatter plot depicting the budget per capita of Boston neighborhoods as a function of the percentage of number of children in the neighborhood (0-17). Note: x-axis values represent percentage out of 1 (0.1 = 10%).



Challenges and Limitations:

Base Project

For the base project, we encountered concerns regarding model complexity, influence of external factors and data formatting to accurately predict budget trends regarding the key questions.

One notable challenge was predicting budget trends accurately when considering the impact of numerous unpredictable external factors. Economic conditions have fluctuated heavily in the 4-year period we analyzed and they play a significant role in shaping budgets. Furthermore, political decisions add another layer of complexity as policy shifts have profound effects on budget allocations and priorities. These two conditions made it challenging to anticipate the future budget trends for FY24 and future years accurately.

One limitation for the base project was creating data visualizations that effectively captured all relevant variables without becoming overly complex, which was a delicate balancing act. Finding the right balance between simplicity and comprehensiveness was difficult, but it necessitated a thoughtful approach to variable selection and model design. For example, we employed the strategy of depicting Top 3 and Bottom 3 for departments which received funding to ensure our visuals are both accurate and practical for understanding the analysis.

Finally, we had to assume consistency in data formatting from the dataset we were given for the base project. This was vital for accurately predicting and forecasting future budget spending patterns across departments and categories. Inconsistent data formats could have introduced errors and resulted in inconsistent spending predictions.

Extension Project

For the extension project, we wanted to ensure the accuracy and reliability of our analysis is essential by addressing various concerns related to data synchronization, temporal gaps, and missing data.

One notable challenge was ensuring that the population and spending data were synchronized for the same years to ensure accurate analysis since mismatched data years could introduce inaccuracies and potentially skew our findings, making it more difficult to draw meaningful conclusions. To address this concern we focused on accessing a synchronized dataset for population per neighborhood and spending per neighborhood for the same years.

The main limitation for the extension project was that the analysis provided was based on census data which is taken every 10 years. Relying solely on census data may result in an incomplete and potentially outdated picture of the population and spending dynamics. This means there could be a loss in client confidence as stakeholders may be concerned about the relevance and

timeliness of the conclusions drawn from the analysis. Furthermore, there is a possibility of economic and social changes which may result in an non-comprehensive understanding of the data and its implications. For example, since the census data was from 2020, the effects of the Covid-19 pandemic may be factored into our analysis. However, at the current time of our investigation, most of the ramifications of the lockdowns and pandemic may not be as influential in population and spending data. Finally, the temporal gap in data analysis could result in missing important trends that have emerged during the years since the census. This can hinder the ability to make accurately informed decisions based on the current conditions. It was difficult to address this limitation as the only census data we had access to, via the City of Boston website, was collected on a 10-year basis.

The last challenge encountered was regarding missing data in the dataset. In order to counter this, we employed data interpolation and cleaning strategies to enhance the completeness and accuracy of our dataset.

Conclusion

Base Project

During the exploratory data analysis, we found that the Execution of Courts Department and the Annual Audit costs reported a significant increase in funding in recent years. We hypothesized that these two are correlated and this increase in funding was done intentionally. We believe that funding increased to these two specific departments because they can work together to bring in more revenue in future years through tax auditing.

Through our visualizations of the data, we found that the budget for the past several years and the project budget for FY24 is indeed increasing. This information can be found in Figure 1, which demonstrates that the various departments of the city of Boston are receiving more funding. This can be due to several factors, such as inflation reducing the value of capital being allocated, and optimism that allocating funds can aid the city to bring in more revenue through tax collecting methods.

In addition, by utilizing data provided by the Spark! team, we were able to identify areas of Boston that received more funding than others. This information can be located in Figure 6, which shows areas such as Charlestown, Downtown, and Chinatown receiving the most funding.

Furthermore, we identified several of the departments that received the most funding. Departments such as the Boston Public Schools department, the Police department, and the Fire department are all among the highest funded departments in the city of Boston. This information can be located in Figure 2.

Extension Project

Our hypothesis stated there could be a relationship between percentage of older individuals and white individuals in a neighborhood and the corresponding allocation of the city budget per capita, presented mixed findings from our analysis and data visualizations.

While our initial belief in a correlation with the population of children (0-17 years old) and higher budget allocation per-capita proved unfounded based on Figure 18, our exploration of race demographics revealed nuanced patterns. Figures 15, 16 and 17 displayed varying correlations between budget per-capita and the racial composition of neighborhoods. There was a noteworthy increase in budget per-capita in areas where the population of white individuals exceeded 65% of the neighborhood seen in Figure 15, however, there was not enough evidence of a strong correlation. Additionally, Figures 16 and 17 highlighted the disparities in budget funding per-capita for neighborhoods with larger populations of African-American and Latin individuals. Figure 16 depicted that neighborhoods with larger African-American populations are among the lowest funded. Figure 17 showed a similar trend for neighborhoods with Latino populations, although there is only one neighborhood with a Latino population larger than 30%, so the correlation is not as pronounced as with African-American populations.

It is also important to mention that Charlestown, Chinatown and Downtown stood out as outliers with the highest budget spending per-capita, while Fenway, Mission Hill and Hyde Park received the least. Despite the high presence of college students in Fenway and Mission Hill, these neighborhoods did not receive commensurate budget allocations. These findings allowed for a deeper exploration of the factors influencing budget distribution across neighborhoods in Boston, challenging assumptions and paving the way for a more nuanced understanding of the complex interplay between demographics and budgetary decisions.

Lastly, we wanted to acknowledge the conclusions we drew from our extension project analysis is limited due to the fact that our insights are based on 2020 census data. We made the assumption that there would be a minimal change in racial demographics and population densities for Boston neighborhoods since 2020. We believe that census data would be collected more frequently if there were more notable changes in these factors in neighborhoods in a shorter period of time. Therefore, we assume to the best of our ability we have captured accurate trends surrounding budget spending per-capita in Boston neighborhoods regarding population density and demographics.

Contributions of Team Members:

Toby: Toby worked primarily on data exploration and analysis. He cleaned and parsed the data for the extension project, merged the new dataset containing population statistics with the original dataset containing project budget information, and created various visualizations to explore that data. He drew some initial conclusions from those graphs that were later refined by other team members for the deliverable report.

Lucy: Lucy performed data analysis on spending by neighborhood, which was combined with the demographic data analysis in the extension project. She also helped write, format, and edit the final report. She met with the TA every week during lab to ensure that the team was on track for the final deliverable.

Deep: Deep worked on collecting/finding data relevant to the extension project and presenting some initial findings from this data relating to neighborhood demographics. Deep also worked on creating some initial representations of the initial data we were given by the Spark! team. Deep also worked on the deliverable reports and presentations.

Thian: Thian worked on all revising and updating the analysis, results and visualizations for all the key base questions for the deliverable 5 report. He also went in to meet the TA to ensure that the requirements for the report would be met. Thian worked on creating the structure for the extension project in the deliverable report by adding the questions and the key takeaways from the findings. He also added the challenges and limitations aspect to the deliverable 5 report.