

CS 506 - Data Science Fundamentals

City of Boston: City Budget Team A

Group Members: Toby Ueno, Deep Patel, Lucia Kisova, Thian Amarasekera

Problem Statement: For this deliverable, we have been working on going through the data we have been given. We have been given data regarding the department and category spending for the city of Boston. We are working on finding out how money has been spent by these different departments for the years 2021, 2022, 2023, 2024. Also, we have been given the data for the projected spending for the 2024 fiscal year. We are aiming to find how money has been spent and allocated for these years.

Data Collection/Cleaning:

Our data was pre-collected from the City of Boston website for the Operating Budget Data; no extra steps for collection were necessary beyond downloading the associated CSV files. For the extension project we have decided on for now, we have found a dataset from the city of Boston regarding how many people live in a particular neighborhood in the neighborhoods of Boston.

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.

Hypothesis: One interesting thing to note is that the Execution of Courts department and Annual Audit costs reported a significant percent increase in funding. 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?

On average, the city budget increased by **5.59%** per year. Budget had a **6.56%** increase from 2021-2022, a **3.56%** increase from 2022-2023, and a projected **6.65%** increase from 2023-2024.

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

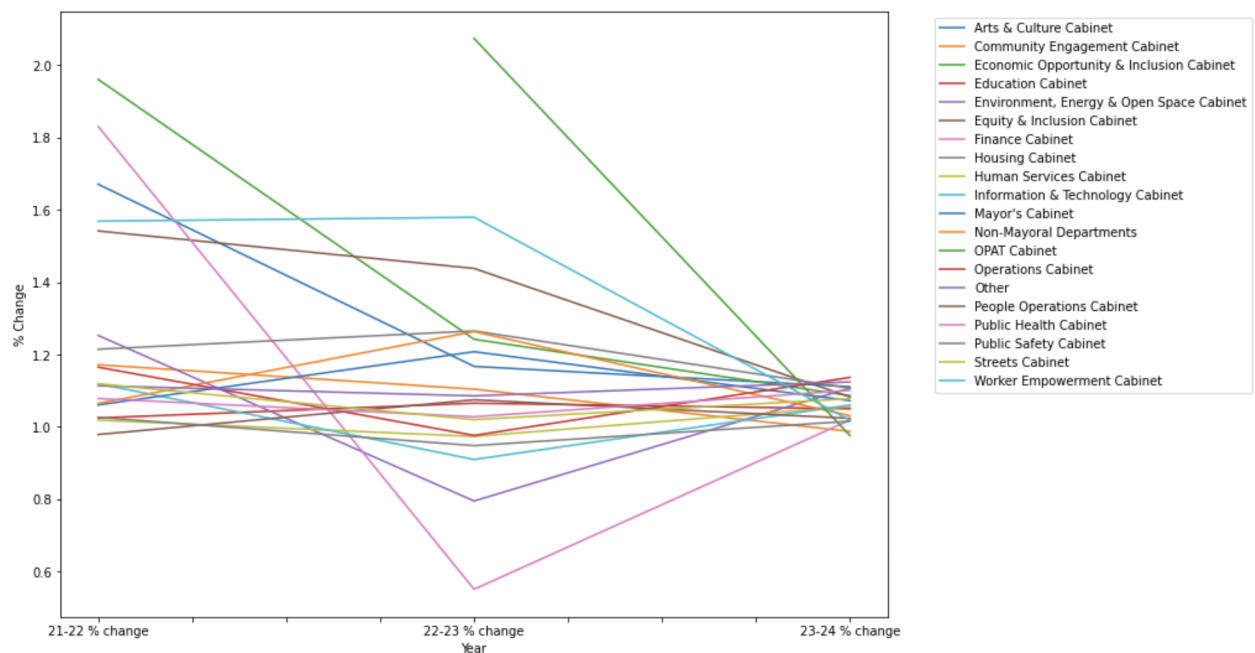
Notably, this excludes the OPAT cabinet, which received no funding in 2021 (it was conceived in 2022).

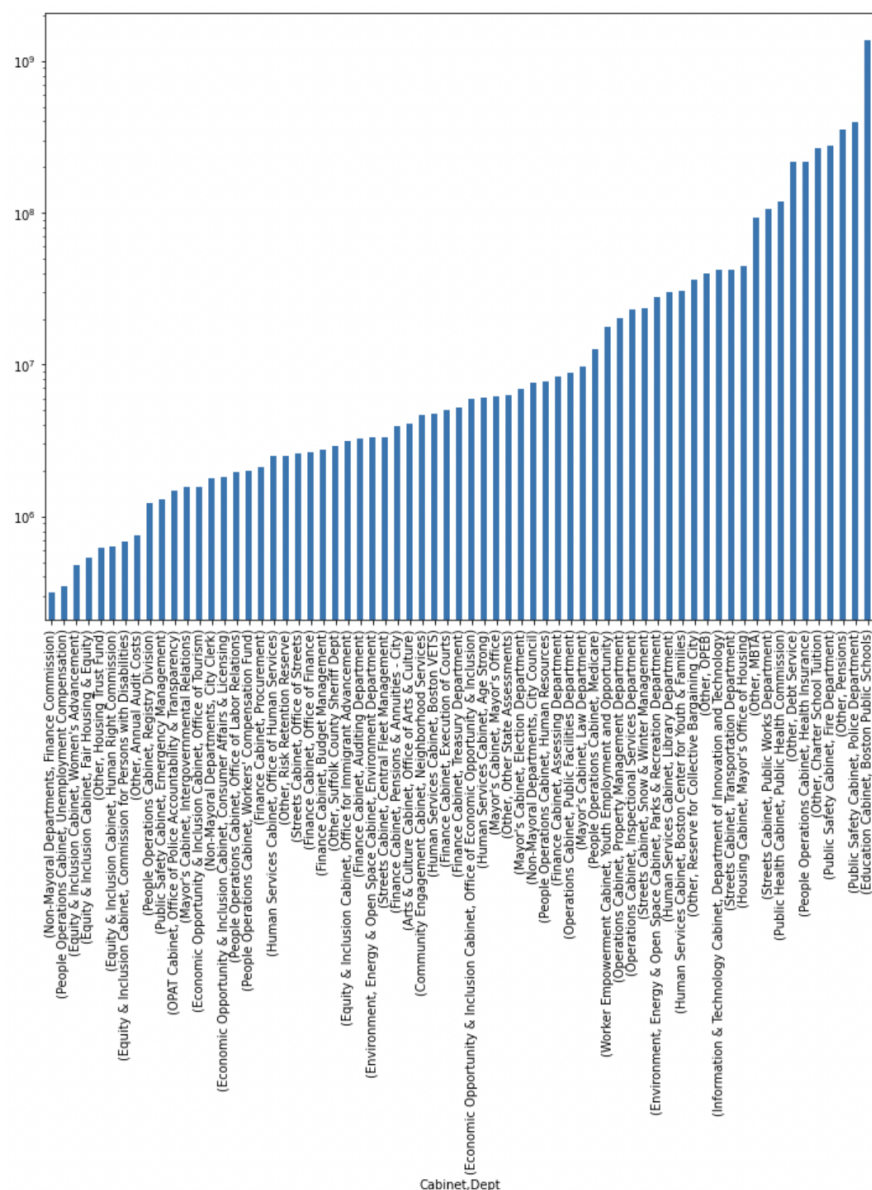
The three departments which lost the largest percent-decrease in funding per year over this four-year period were:

- 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

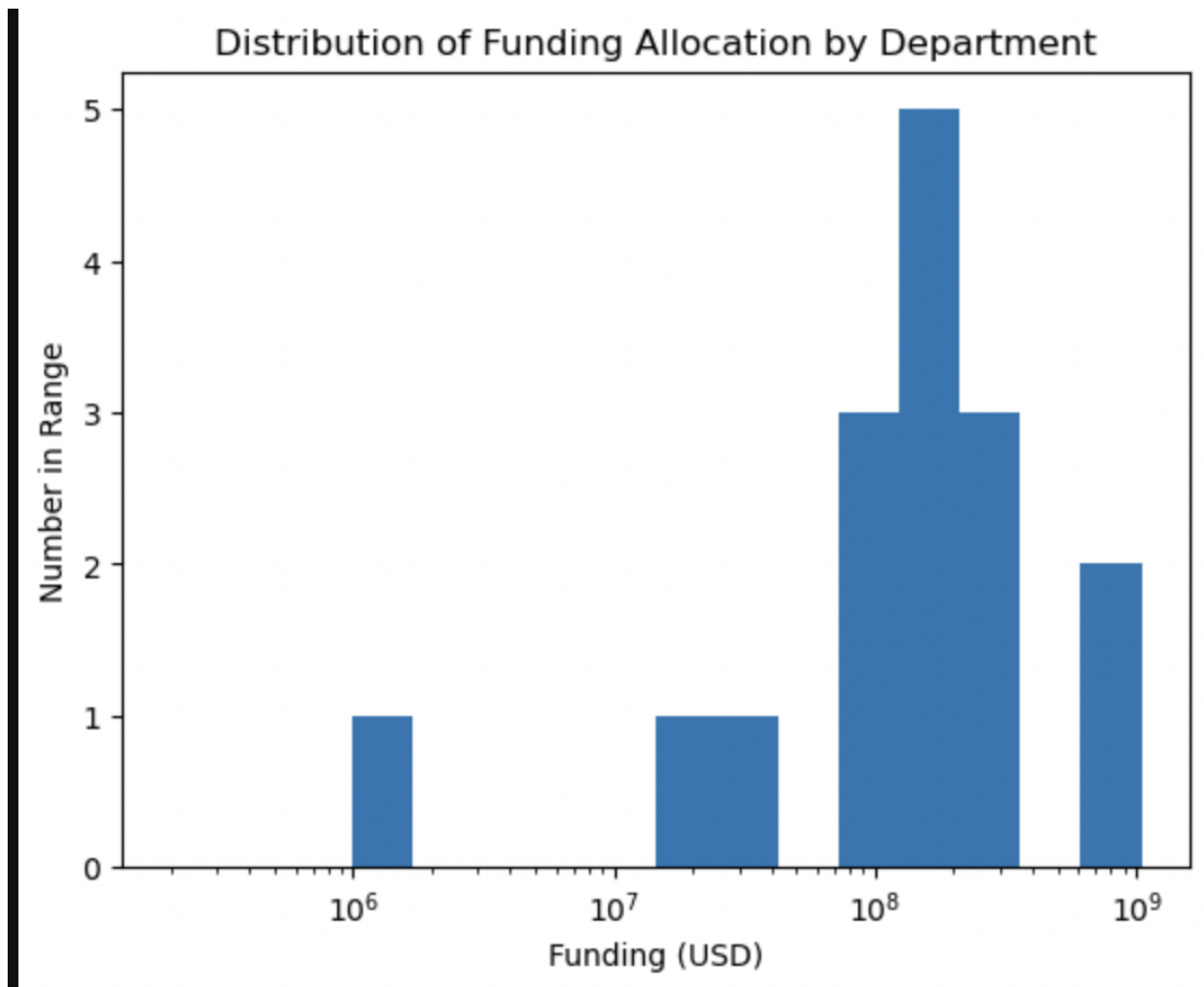
The Police Department and the Central Fleet Management were the only other two departments to report an average yearly decrease in funding over this period. All other 61 departments reported an increase.

Visual: The data given shows this increase as well. A general increase year over year can be seen by looking at the percentage increase from each year. This data shows that while there have been some outliers, most departments show an increase in spending.





Visual: A similar bar chart shows the distribution of funding for the year 2024 in terms of how many departments receive that much funding. As seen, many departments are receiving a large amount of funding.



Key Question #2: How has spending per program changed over 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

The Trust project is somewhat of an outlier, since its budget in 2021 was relatively low (\$631.62).

Again, this excludes the OPAT cabinet, which received no funding in 2021 (it was conceived in 2022).

The three programs which lost 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

151 projects (82.07%) received an increase in funding, whereas 33 (17.93%) received a decrease.

Contributions of Team Members:

Toby: Data processing and cleaning. Toby has gone through the given data from the city of Boston and looked at the format of the data. He found how the data was formatted and created some visuals.

Lucy: Lucy met with TPM on 10/19 to get pointers for how to start data analysis (notes in scrum report docs). She was working on data analysis for the proposed capital budget, comparing the total proposed budget by department for the upcoming fiscal year. Lucy also did cleaning and some initial visualizations for this second dataset.

Deep: Deep viewed both of the datasets and created some visuals for these datasets. He created graphs so the team could get an initial view of how the data was formatted, structured, and what it represented. Deep found some data that may be helpful in the extension project. Attempted to meet with Tpm during office hours but TPM was not present.

Thian: Thian did some research on the datasets that might be needed. He also created the scrum reports for the group. Thian also helped the group members explore the given data.