CS506 Project Meeting Notes

First Weekly Scrum Report - 10/05/23
We all met and got to know each other and discussed our motivations and the goals for the project. We went through the project documentation and collectively answered the following questions that will help us start with the project.

What is the project focus/overall goal?
The overall goal of this project is to analyze the budget of the City of Boston. In the analysis, we will try and understand how the budget is spread across various categories like departments and geography and showcase how it has changed over time. This analysis aims to provide valuable insights into the city's financial management and empower stakeholders to make better-informed decisions regarding city services and finances.

Why is this project important?
This project is important because it has the potential to have an impact on the lives of millions of people. This data-driven approach not only ensures efficient use of public funds but also engages citizens in the budgetary process, fostering trust in government and promoting informed governance practices.
This is what motivated our team to opt for this project. It will provide transparency and help demand accountability from the budget planners. By analyzing spending patterns across departments, budget categories, geography, and programs, we can make informed decisions, identify disparities, and evaluate program effectiveness.

What type of data will you collect or be analyzing?
We will be analyzing data about the way the city of Boston spends its annual budget, and how this has changed over time. The focus will be on per capita spending by: department, geography, program, and budget category. That is with the goals of analyzing and finding the impact of revenue shifts, socio-economic indicators, and funding sources.
Specifically, we will be collecting and analyzing two primary sets of data pertaining to the City of Boston's budget. The **Operating Budget Data** offers a comprehensive breakdown of annual expenditures, detailing allocations for various departments such as education, police, and housing, and encompassing costs for services and personnel like teachers and firefighters. Meanwhile, the **Capital Budget Data** sheds light on large-scale investments in the city's physical assets. It elaborates on funding sources, from bonds to grants, and provides explicit project descriptions and allocations.
We'll also be exploring the following data sources :
Revenue data - To understand how the revenue shifts impacted different departments, programs, geographies
Economic Data like housing data

Socio-economic data like Social vulnerability index
Departmental budget data to understand the division among departments better

What are the potential limitations of the project?

- Ethical considerations can limit access to certain data especially if it contains private or sensitive information
- There could be inconsistencies among how the budgets are recorded each year and this would impact the trends over time.
- It might be difficult to compare data during the years because of economic inflation/fluctuations
- Trends across different areas like operating vs city budget and various different departments might not emerge clearly through visualizations if they are haphazardly changing over the years

What are your next steps? Divide tasks amongst the team
The next steps are to explore the various data sources that we have available and identify how we can use them to work towards the project goals stated above. We will also be looking to identify other potential sources of data through help from the PM and other SME's that we can get access to.
The task for all team members is to get familiar with the data for the main problem and also explore the potential of at least one of the extension projects and determine if it is feasible.

10/06/23
Attended the lab and shared our progress. Naima and Jason also attended.
Created a discord server - which will act as our official medium of communication from now on.
Created tasks on the trello board for all members for this week.
Download and explore the data for the base project. Links are in the discord and trello board.
Please join the discord and trello.
Next meeting is on Monday after DS class. Please complete the trello tasks by then.

Weekly Scrum Report - 1 10/07/23 - 10/13/23

WHAT WE WORKED ON THIS WEEK

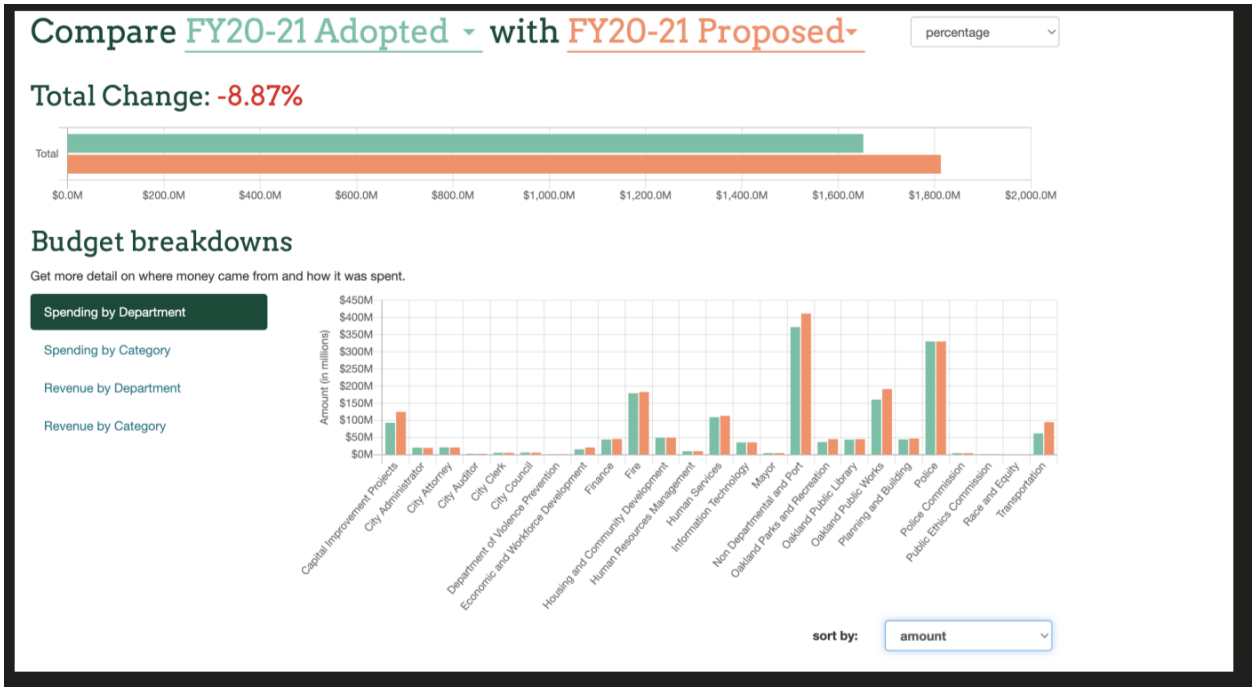
We had 2 team meetings - on 10/11/23 and 10/13/23

10/11/23
Members in attendance - Gaurav, Jason, Haokun, Alex

Everyone gave their input based on their understanding of the problem statement and their exploration of the data.

I took inspiration from the official budget websites of Arlington, Oakland and Philadelphia and proposed that for our use case we can do visualisations similar to the following 4 graphs.

1. Histogram

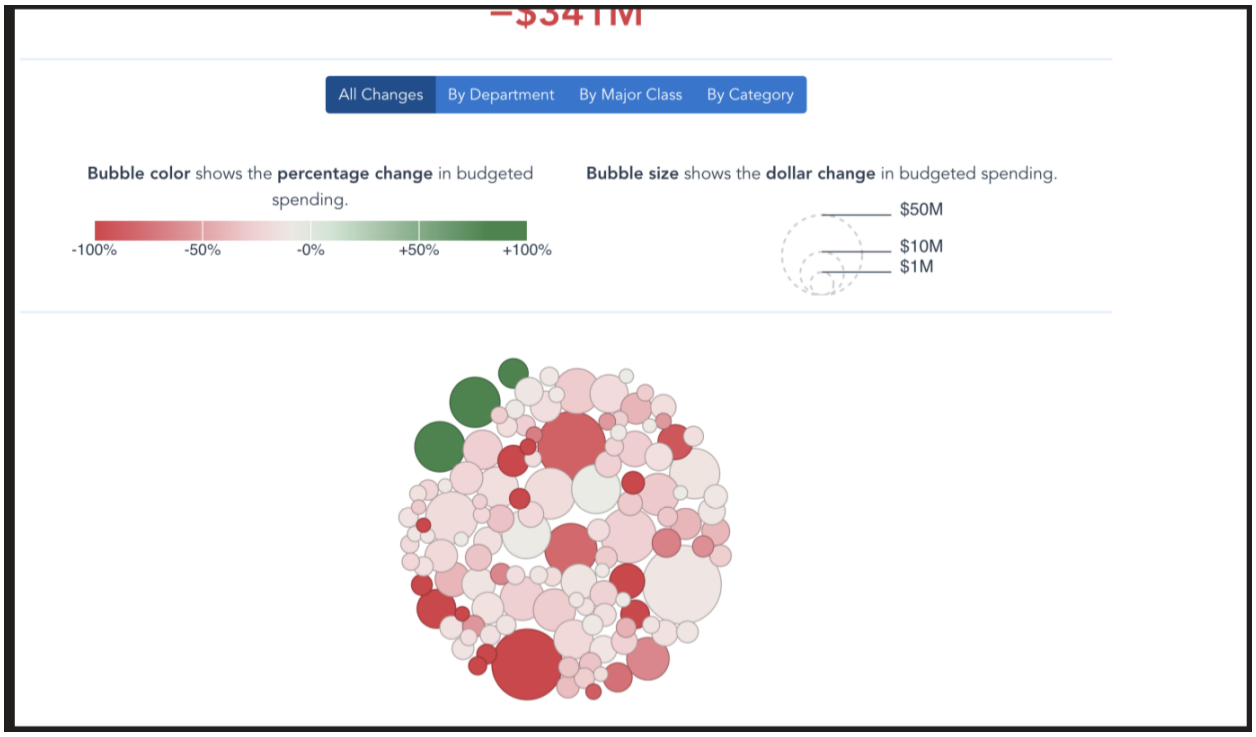


The histogram will be suitable to show the distribution across different - in fields which have around 10 columns like Expense Category in Capital Budget Dataset

```
df.ExpenseCategory.unique()
```

```
array(['Personnel Services', 'Contractual Services',
       'Supplies & Materials', 'Current Charges & Obligations',
       'Equipment', 'Other Expenses', 'Fixed Expenses'], dtype=object)
```

2. Bubble Graphs



This graph is suitable to show the distribution across the attributes where the number of unique values is very large like the ‘Program’ attribute in Capital Budget dataset - which has 220 unique values. Through the colours we can also show if the allocation is increasing/decreasing through the years.

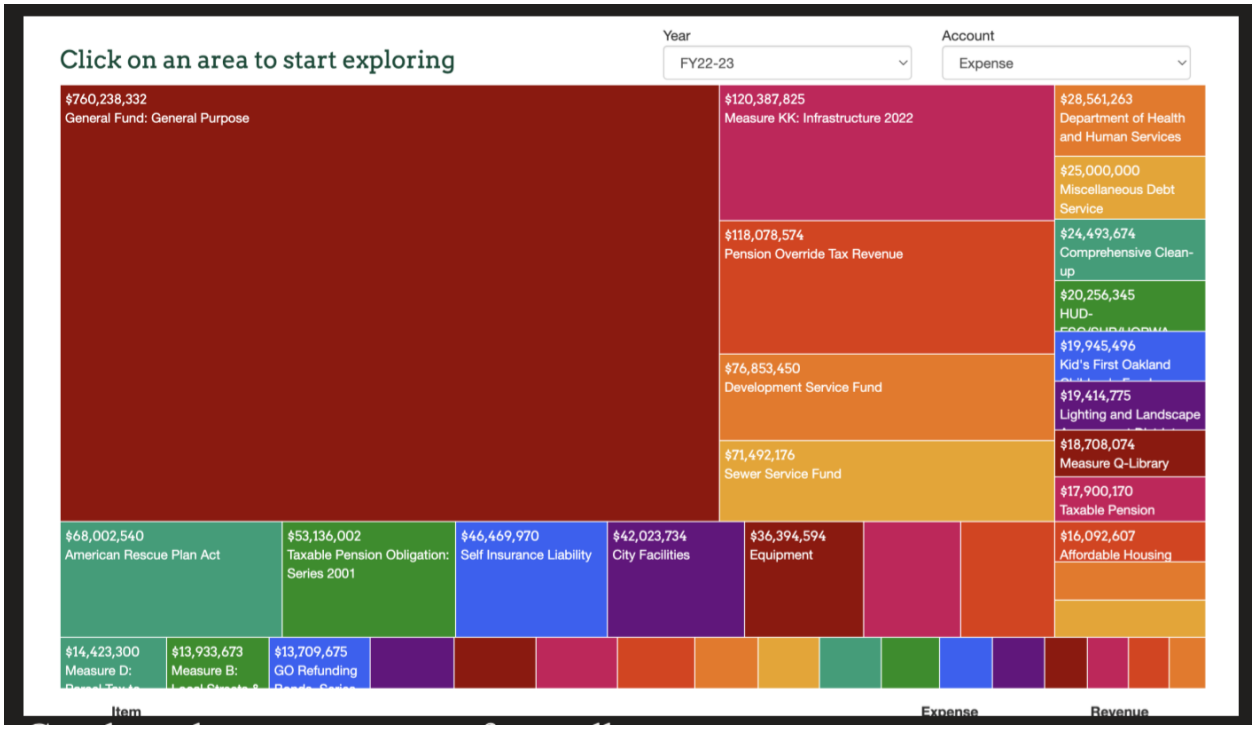
```
len(df.Program.unique())
✓ 0.0s
220
```

Cabinet - bar chart - showing distribution

```
df.Program.unique()
```

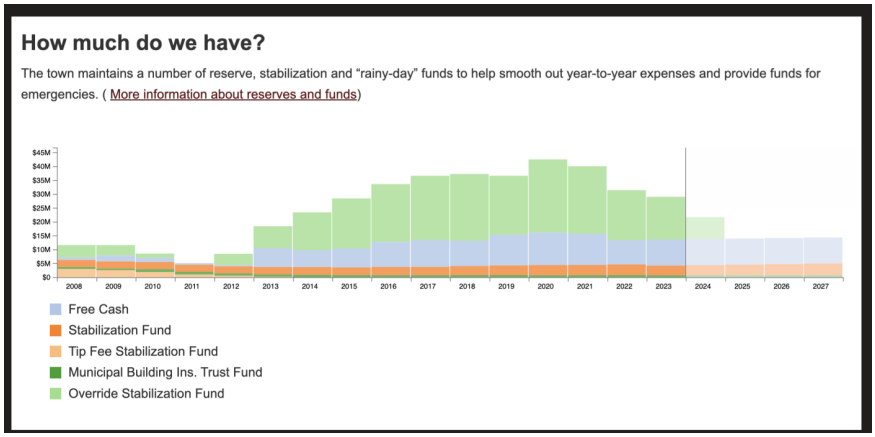
```
array(["Mayor's Administration", "Mayor's Executive",
       "Mayor's Policy & Planning", 'New Urban Mechanics',
       "Mayor's Communications", 'Election Division', 'Listing Board',
       'IGR', 'Grants Administration', 'Law Operations', 'Litigation',
       'Government Services', 'Equity', 'Equity & Inclusion', 'Diversity',
       'MORRE', 'LCA', 'Human Rights', 'Immigrant Advancement',
       "Women's Commission Program", 'BMA', 'Fair Housing Commission',
       'Human Rights History', 'LGBTQ+ Admin', 'Disabilities',
       'OPAT Commission', 'Civilian Review Board (CRB)',
       'Internal Affairs Oversight Panel (IOAP)', 'PM Administration',
       'Building Operations', 'Alterations & Repair', 'Enforcement',
       'Security Systems', 'Animal Control History', 'Building Systems',
       'PFD Capital Construction', "ISD Commissioner's Office",
       'ISD Administration & Finance', 'Buildings & Structures',
       'Field Services', 'ONS Administration', 'Neighborhood Serv',
       'Office of Civic Organizing', 'Boston 311', 'Arts and Culture',
       'Economic Development', 'Equity & Inclusion History',
       'Small & Local Business', 'Consumer Licensing', 'Consumer Affairs',
       'Licensing_Board', 'Supplier Diversity Administation',
       'Tourism Administration', 'Film & Special Events', 'Tourism',
       'Boston Residents Job Policy', 'Worker Empowerment Administration',
       'Youth Employment and Opportunity Admin', 'Alternative',
       'Early Learning Center', 'Elementary', 'Exam', 'High', 'K-12',
       'K-8', 'Middle', 'Middle/High', 'Special Ed, K-12',
       'Superintendent', 'Student Support Services (C)', 'Strategy',
...
       'Finance Commission Program', 'Pensions', 'Debt Service',
       'Charter School Tuition', 'MBTA', 'Other State Assessments',
       'Suffolk County Sheriff Dept', 'Reserve', 'Annual Audit Costs',
       'Housing Trust Fund', 'Reserve for Collective Bargaining City',
       'OPEB', 'Risk Retention Reserve', 'Tax Title'], dtype=object)
```

3. Heat Map



Heat Map will be useful in categories like Department where there is a stark difference between the distribution across ‘different departments’. The area can represent the budget allocated to the department and the colors can be used to represent another dimension.

4. Stacked Bar Chart



This can be used to show the picture of division of budget across various types of an attribute and sub-division within it.

In the meeting we finalized the 4 graphs. We planned to work on it in the next week.

10/13/23
Members in attendance - Gaurav, Naima

WHERE DO WE NEED HELP
Asked PM about the discrepancy in the Departments in both the data sets - Capital and Operational Budget. Capital budget also has 2 department columns. PM informed that a separate slack channel will soon be set up where we can ask questions to the clients.

TASKS FOR NEXT WEEK

1. Perform data cleaning across all 3 data-sets
2. Finalise data cleaning methods that we'll use across the project so that there is no discrepancy
3. Start making visualisations and discuss wherever we fave any issue.

TASKS ALLOCATION for next week
Gaurav - Data Cleaning for Operating Budget
Naima - Data Cleaning for Capital Budget
Sophia - Data Cleaning for Revenue
Alex - Capital Budget Graphs - Histogram, Stacked Bar chart
Jason - Operating Budget Graphs - Histogram, Stacked Bar chart
Haokun - Revenue Grpahs - Histogram, Stacked Bar chart

Dataset descriptions (Naima)

Dataset 1: FY24 RECOMMENDED CAPITAL BUDGET PLAN
https://data.boston.gov/dataset/capital-budget/resource/c62d666e-27ea-4c03-9cb1-d3a81a1fb641

The provided dataset details Boston's FY24 Recommended Capital Budget Plan, outlining various infrastructure projects, their statuses, scopes, and budget allocations. Key columns include the department responsible, project name, scope of work, current status, and budget details, such as authorizations, grants, and total budget. For instance, the dataset includes projects like BCYF Security and Technology Upgrades with a total budget of $2,000,000, and BCYF North End Community Center with an $88,000,000 budget, indicating the city's diverse investment in infrastructure developments.

Brief descriptions of the columns:

1. Department: Indicates the city department managing the project.
2. Project Name: The title or name of the specific infrastructure project.
3. Scope Of Work: A brief outline of the project's objectives and tasks.

4. PM Department: Department overseeing project management.
5. Project Status: The current status or phase of the project.
6. Neighborhood: The city area or neighborhood where the project is located.
7. Authorization Existing/FY/Future: Funds allocated currently, in the fiscal year, and anticipated in the future.
8. Grant Existing/FY/Future: Grant funds allocated currently, in the fiscal year, and anticipated future grants.
9. GO Expended: The amount of Government Obligation funds already spent.
10. Capital Year 0/1/25: Projected capital expenditures for the current, next, and 25 years ahead.
11. Grant Year 0/1/25: Projected grant expenditures for the current, next, and 25 years ahead.
12. External Funds: Funding from non-city budget sources.
13. Total Project Budget: The combined total of all funding sources for the project.

Dataset 2: FY24 RECOMMENDED OPERATING BUDGET
https://data.boston.gov/dataset/operating-budget/resource/8f2971f0-7a0d-401d-8376-0289e3b810ba

The dataset outlines the FY24 Recommended Operating Budget for the city of Boston, totaling $4.28 billion, a 6.8% increase from FY23. It contains detailed records of projected expenses for various city departments and programs, derived from the City's General Fund, for the fiscal year starting July 1, 2023, and ending June 30, 2024. Each row represents a unique combination of department, program, and expense category, showcasing a clear breakdown of where and how funds are allocated and spent. The data allows for a detailed analysis of the financial planning and allocations across different city departments and services.

Brief descriptions of each column in the dataset:

1. Cabinet: The top-level organizational unit, often associated with city governance.
2. Dept: The specific department within the cabinet responsible for a program or service.
3. Program: Individual programs or initiatives managed by the department.
4. Expense Category: The type of expense, such as personnel services, contractual services, etc.
5. FY21 Actual Expense: The actual expenses incurred during Fiscal Year 2021.
6. FY22 Actual Expense: The actual expenses incurred during Fiscal Year 2022.
7. FY23 Appropriation: The budget allocated for expenses in Fiscal Year 2023.
8. FY24 Recommended: The proposed budget for expenses in Fiscal Year 2024.

Dataset 3: FY24 ADOPTED REVENUE BUDGET
https://data.boston.gov/dataset/revenue-budget/resource/9c58a0c8-02c3-4712-850b-7b06f16e8fd5

This dataset provides information on the budget adopted for FY24 by a city and how it compares to the budget and actual revenues of previous fiscal years (FY21, FY22, FY23). It includes details of both recurring and non-recurring revenues and breaks them down into different categories and accounts.

A brief description of each column in the dataset:

1. Revenue Category: Describes the type of revenue, such as Property Tax Levy.
2. Account: Specifies the particular account under each revenue category, like Real Estate Taxes, Personal Property Tax, etc.
3. Cabinet: Indicates the higher-level organizational entity or section of the city government overseeing the revenue accounts, in this case, Finance.
4. Department: The specific department within the cabinet that is responsible for the account, such as the Assessing Department.
5. FY21 Actual: The actual revenue collected under each account during Fiscal Year 2021.
6. FY22 Actual: - The actual revenue collected under each account during Fiscal Year 2022.
7. FY23 Appropriation: The anticipated or budgeted revenue for each account for Fiscal Year 23; it may not be the actual collected amount.
8. FY24 Adopted: The adopted or finalized budget for each account for Fiscal Year 24, indicating planned or expected revenues.

Each row in the dataset provides detailed information for a specific account, including the actual revenues from past fiscal years, the appropriated budget for FY23, and the adopted budget for FY24, organized by the responsible cabinet and department.