Fall 2022 Councilor Mejia x City Services Project

Team 3 - Deliverable 1

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 - Team Lead	ddumouch@bu.edu
Zihao Shen	zhaojun@bu.edu
Tian Tan	tiant@bu.edu
Lu Yao	ly21@bu.edu

● Collect and pre-process a preliminary batch of data

In this step, we filtered and preprocessed the dataset. We dropped the dataset, 'Dataverse updated census data', which was created by Harvard Dataverse because it only followed up to 2019.

We decided to use the 'Census dataset for Boston', which contains census data for 2020 and it is an official government release. This dataset actually contains four sub-datasets, namely 'Census Tract', 'Census Block Group', 'Boston Voting District' and 'Boston Neighborhood'.

To be more explicit, currently, we use the 'Boston Neighborhood' dataset and the 'Business Assistance Funds' dataset published by the Boston government.

● Perform a preliminary analysis of the data

'Boston Neighborhood' Dataset: It counts the race composition and number of people in each neighborhood block in the Boston area. At the same time, it also records the types of houses in each neighborhood block, the size and the number of householders. It also divides statistics by age.

'Business Assistance Funds' dataset: It records recipients of funds in each neighborhood block in the Boston area. From the bar chart, we can see there are most companies which request funds in Dorchester, followed by Back Bay. Very few small businesses request funds in the Downtown/ financial District.

● Answer one key question

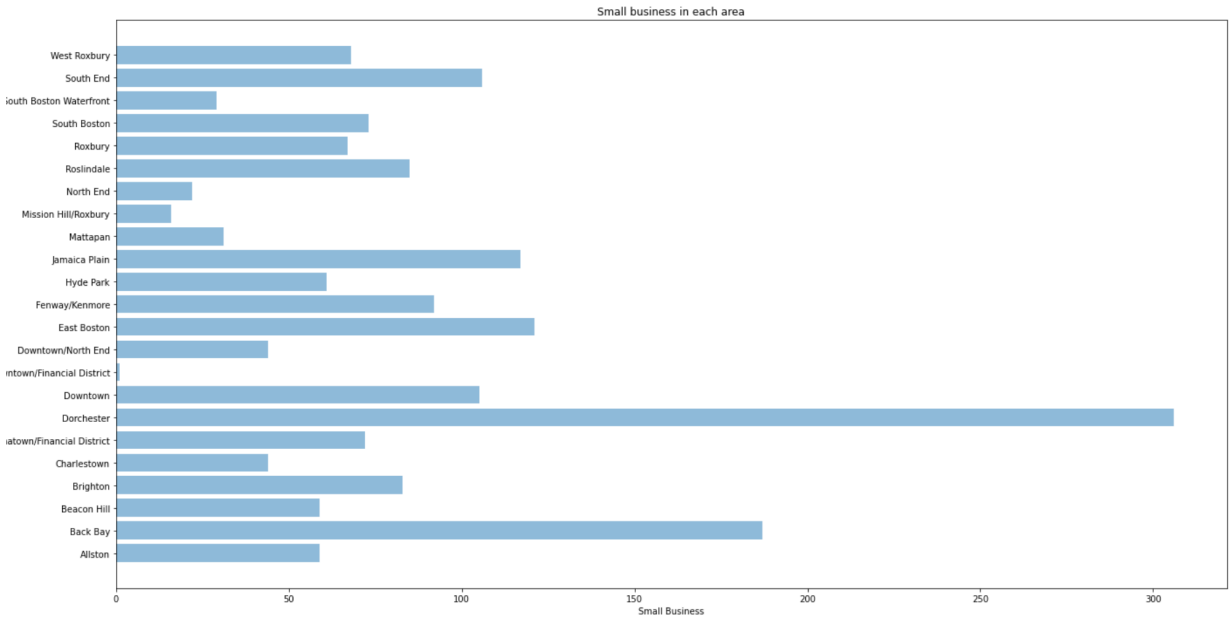
Where did business assistance go during the pandemic?

From our analysis of the funding dataset, we got the following graph. We can find that Dorchester has the most fund recipient companies, with more than 300, followed by Back Bay, and Downtown/Financial District has the fewest fund recipient companies, close to 0.

Thus, in general, more funds/business assistance flow to the 'Dorchester' neighborhood block.

Fall 2022 Councilor Mejia x City Services Project

Team 3 - Deliverable 1



● Refine project scope and list of limitations with data and potential risks of achieving project goal

Scope: 1) Track and analyze the flow of funds. 2) Analyze the demographic data of the flow of funds to the region. 3) Combining the previous two to analyze the allocation of funds.

Limitations: 1) The relationship between the various datasets is relatively vague, and no specific connection can be found. 2) The datasets are complex and not intuitive and concise. A dataset has many sub-datasets. 3) Census datasets do not have specific latitude and longitude coordinates, making it difficult to draw on a map for visualization.

Potential Risks:

We are missing a specific amount of funds for each neighborhood block. Only the number of fund recipient companies within each neighborhood block is known, but this does not accurately reflect the specific funding for each neighborhood block.

Furthermore, even knowing the exact amount of funds, we lack the true value of where the funds were expected to be allocated.

At the same time, we do not know what specific relationship we need to establish between the flow of funds and the distribution of census composition.