

## **Councilor Worrell x Small Business Final Deliverable**

Zula Nyamtur, Thomas Lo Bello, Tiffany Chen, Lesley Chen  
Fall 2022

### **Background and Motivation**

The purpose of this project is to understand the current landscape of small businesses in District 4. We want to understand the current landscape of small businesses in District 4 and find features that will attract small businesses to the area while understanding the current issues small businesses face. District 4 is one of the most diverse districts and historically under-resourced in Boston, and current and previous councilors have been working on advocating for this district.

An analysis of the business environment will offer insight into the successes and shortcomings of small businesses in District 4. With this information, District 4 Councillor Brian Worrell can make informed decisions on what aspects to improve upon and focus on to most effectively support small businesses, as well as make District 4 an attractive place to locate to. Allowing small businesses to thrive can improve the community as a whole. With the growing population of chain stores, we are interested in seeing how small businesses are faring in comparison and ways for them to succeed in the community. The data used include the BPDA data, District 4 shapefile, as well as the shape files of the other districts, data from the Certified Business Directory of certified businesses, and Employment Data of Boston neighborhoods.

Based on our preliminary analysis described below, we decided to define a “small business” as businesses with fewer employees and/or less annual revenue than a regular-sized business or corporation and limited this to approximately 50 employees. In addition, we determined that this does not include chain businesses with few employees, as chain businesses have select advantages in that the central organization assumes responsibility for the finances of its units.

### **Data collection and Exploration**

#### Preliminary Analysis:

In our initial exploration of the businesses in Boston dataset, we noted that Estimated\_Employment may be a relevant feature to study. Estimated\_Employment is the number of employees a business has the resources for, and this may be indicative of the size of the businesses and its success. We filtered out the businesses for each district and discovered that in District 4, most businesses are in the employment range of anywhere from 2 - 10 employees with some operated by the owner only. However, we found one outlier, Fields Corner Store (located just on the border of District 4 and District 3) with 937 employees. This may be an inaccuracy in the dataset or this means Fields Corner Store is too large to be considered a “small business”.

In comparison with District 6, a neighboring district to some parts of District 4, District 4 lacks the middle ground employment rate that District 6 has. District 6 not only has more business in the 2-10 range, but also in the 10-30 range as well. The number of businesses in the 2-10 range is essentially doubled from D4 to D6. There is also a larger number of non-small businesses based on the estimated\_employment. There may be factors such as district size that contribute to a larger number of businesses in D6 or a larger degree of success, which we will explore in the extension project.

We looked at District 7, which has a similar demographic to District 4, and observed that District 7 also outweighs District 4 in a similar way that District 6 does. There is a healthier middle ground in District 7 of businesses with 10-30 estimated\_employment, though not as drastic as District 6. The number of businesses in the 2-10 range is actually very similar to District 4, but differences are apparent/show beyond 10 employees.

The number of businesses in District 4 is 411; based on our definition of small businesses, we found that the number of small businesses in District 4 is 406. This preliminary study of the estimated\_employment of businesses in District 4, which is indicative of the successes and resources available of businesses in District 4, informs our goal of finding reasons for the failures and successes in District 4.

2.0	71
3.0	68
5.0	44
4.0	40
0.0	38
1.0	27
8.0	22
9.0	20
7.0	16
6.0	15
10.0	14
12.0	6
11.0	4
50.0	3
25.0	3
13.0	2
23.0	2
14.0	2
24.0	1
937.0	1
20.0	1
40.0	1
29.0	1
19.0	1
60.0	1
70.0	1
45.0	1
44.0	1
16.0	1
91.0	1
34.0	1
56.0	1

Figure 0. Left Column: estimated\_employment, Right Column: Number of businesses with that estimated\_employment

## Base Project:

### What businesses exist?

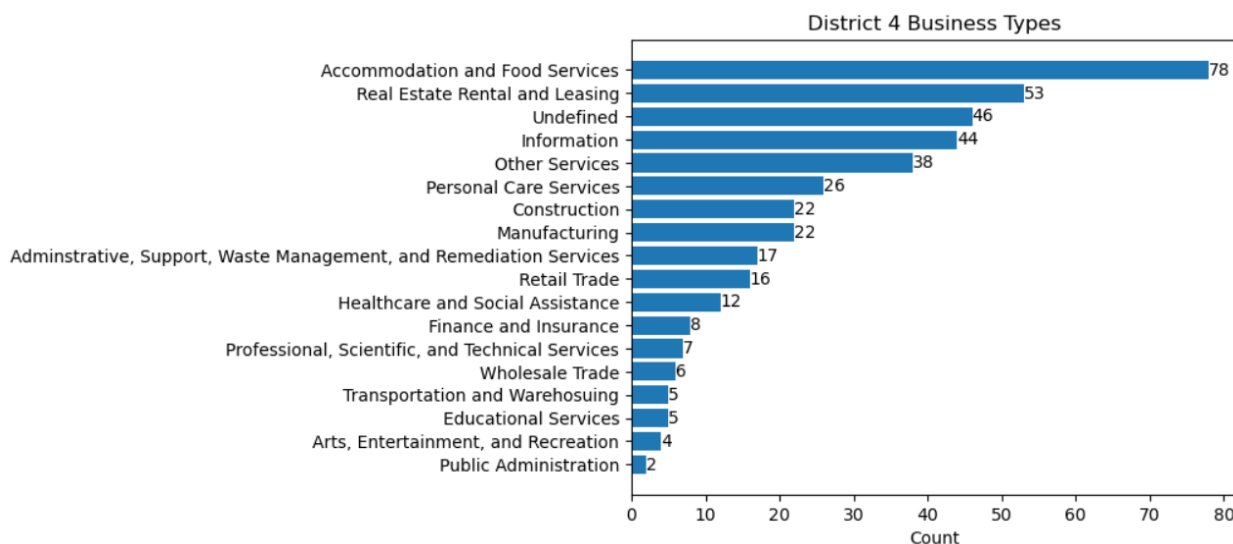


Figure 1. Business Types in District 4

District 4 has a variety of businesses including but not limited to, Full Service Restaurants, Beauty Salons, Financial Transaction Offices. The most prominent business types are accommodation and food services, real estate rental and leasing, information (wireless telecommunication carriers, newspaper publisher, radio stations) and personal care services (Figure 1). Compared to other districts in Boston, District 4 has an overall lower number of businesses. From looking at the two most prominent businesses types in District 4 we can see that District 4 has the second lowest number of these business types out of the 9 districts (Figures 2 & 3).

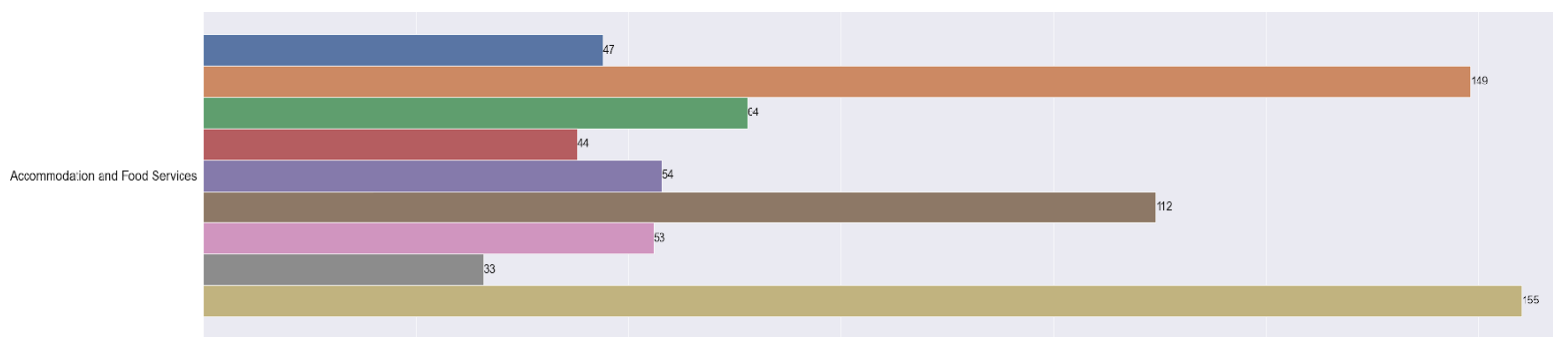


Figure 2. Number of Accommodation and Food Service Businesses Across All Districts

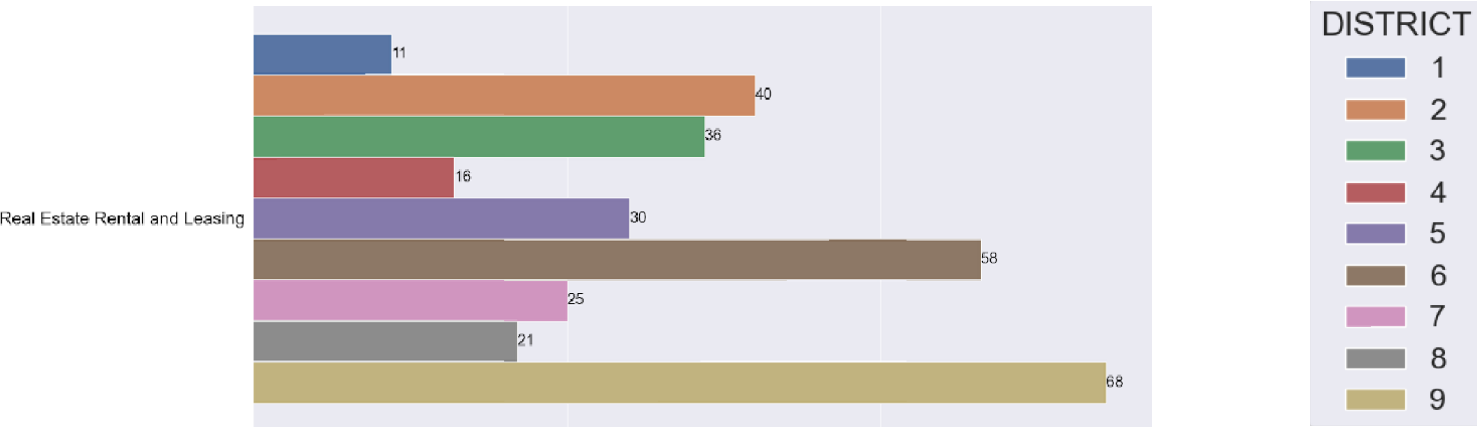


Figure 3. Number of Real Estate Rental and Leasing Across All Districts

What businesses are missing?

Although District 4 covers a large variety of business types, some businesses have plenty of room to grow within the district. There is a limited number of *educational services* in District 4 – a total of 8 businesses fall into this category which includes schools, tutoring, Brazilian jiu jitsu, and yoga. District 1 has a total of 5 educational services which is comprised of driving schools, sports, and family tutoring, having the lowest number of educational services across all districts. District 6 has the highest number of educational services among all districts with 24 businesses consisting of early care and education, martial arts, yoga, art academy, dance, and driving school. Seeing that District 6 neighbors some parts of District 4, it may be the case that residents of District 4 travel to District 6 in pursuit of some of these educational services that are not present within District 4. Therefore District 4 could benefit from an increase in these services to increase engagement and maintain more business within the district.

We also noticed a lack of services in *arts, entertainment, and recreation* – there is one fitness center and one fitness instruction site. However, compared to other districts the gap is not as big as it seems. It is still the case that District 4 can benefit from more services in arts, entertainment, and recreation as it is one of the districts with a lower number of such services.

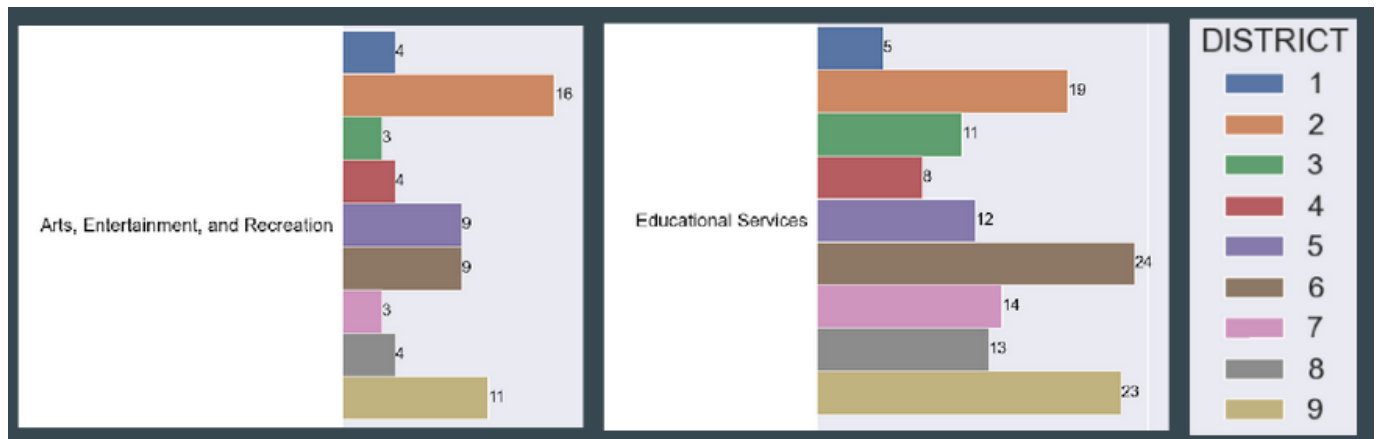


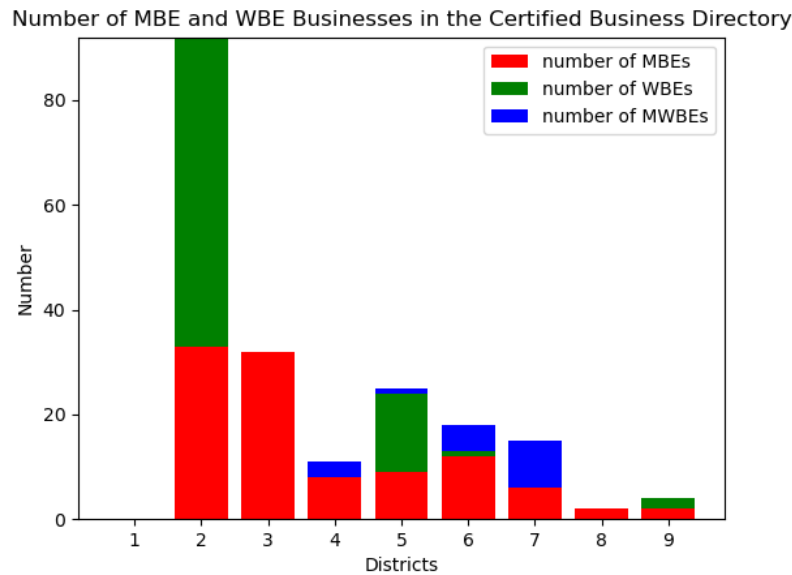
Figure 4. Number of Arts, Entertainment, and Recreation Across Districts

Figure 5. Number of Educational Services Across Districts

#### How many WBE (women owned) or MBE (minority owned) businesses?

Using the [certified businesses directory](#), we looked for businesses in each district's dataset that are certified as MBE and WBE. We merged the certified businesses directory dataset with each district's dataset on 'street\_address' instead of business name due to the possibility of different names or different spelling across different datasets occurring which happened before. However, after we did so, we realized that different companies may share the same address and not all companies at that address may be MBE/WBE/MWBE. Therefore some of the numbers below may be overestimates.

We found that only 11 businesses in District 4 have been registered as MBE. Of those 11 businesses, 3 are also woman-owned. The certified business directory dataset contains 680 data entries of businesses from all over the US. To build a more comprehensive picture of the state of businesses in District 4, we explored the number of businesses in the other districts that have been registered as MBE or women-owned and compared the number between districts. This may reveal information about how accessible it is to be registered, and whether the city council should be advertising these resources to small businesses. Compared to the other districts, District 4 is among the bottom three districts with the smallest number of minority-owned and/or women-owned businesses.



How many of the businesses are chains vs. independent operators?

Due to no existing dataset that contains *all* chain store names, we created our own dataset using information from websites like [Top 100 Retail Chains 2022](#) and [restaurant chains](#). This dataset (chains\_dataset.csv) is not comprehensive at all so we can only perform basic chain store cleaning to get independently operated businesses. Using our created dataset of chain stores, we filtered through District 4 businesses and discovered that 6 businesses are chain stores. This number may be inaccurate (an underestimate) due to the way businesses are named and due to the fact that not all chain stores are accounted for in our created dataset. So 411 total businesses - 6 chain stores = 405 independently operated businesses (an overestimate).

1 isInChains

0.1s

Python

	latitude	longitude	business_name	street_address	zip	NAICS_2017_6digit_code	NAICS_2017_6digit_desc	NAICS_2017_2digit_code	NAICS_2017_2digit_desc
68	42.306190	-71.080272	burger king	100 Washington St	2121	722511	Full-Service Restaurants	72	Accommodations and Food Services
138	42.295882	-71.071859	macy's	450 Washington St	2124	452210	Department Stores	45	Retail and Trade
257	42.298881	-71.062183	target	500 Geneva Ave	2122	452210	Department Stores	45	Retail and Trade
262	42.296989	-71.060642	subway	1576 Dorchester Ave	2122	722511	Full-Service Restaurants	72	Accommodations and Food Services
276	42.299236	-71.060883	mcdonald's	1513 Dorchester Ave	2122	722511	Full-Service Restaurants	72	Accommodations and Food Services
320	42.298881	-71.062183	dollar general	500 Geneva Ave	2122	452319	All Other General Merchandise Stores	45	Retail and Trade

[chain stores in District 4]

As seen in Figure 7 below, the proportion of chain stores as compared to non-chain stores in District 4 doesn't appear to be much different from the other districts.

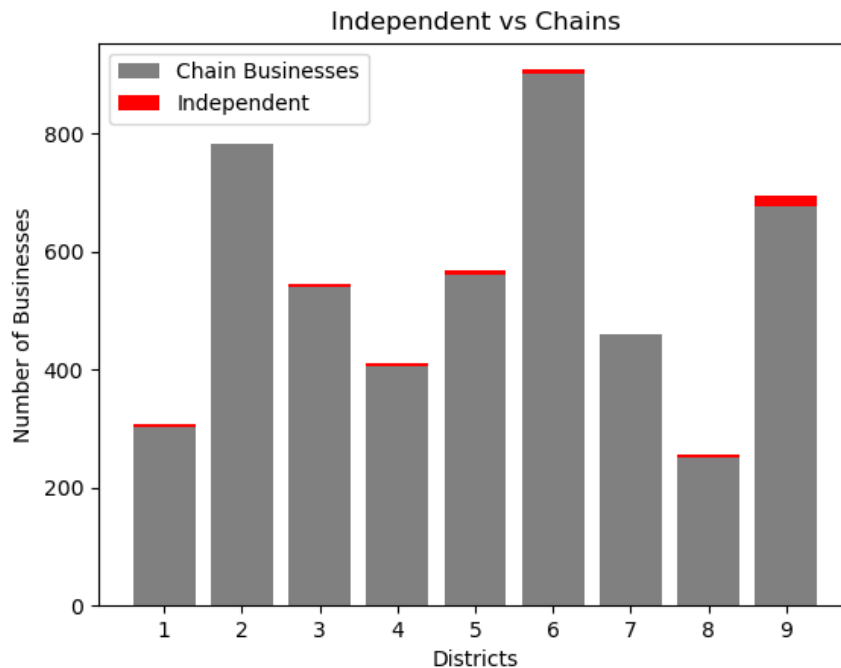


Figure 7. Chain Businesses vs Independent Businesses Across All Districts

This also doesn't take into consideration franchises and whether or not franchises should be considered in the same category as independently owned business/small businesses due to the advantages of being a franchise.

#### Extension Pitch:

Using the employment and spending data, we analyzed the district's monetary status as it relates to small businesses and compared it with other districts to identify areas of improvement for District 4. This extension would analyze the performance of small businesses in District 4. By comparing the performance across businesses, we hope to determine areas of improvement that are likely to aid in the success of a small business and ascertain the business landscape that small business owners entering District 4 can expect to encounter. With this information, our goal is to implement a foundational report on how District 4 can become more appealing to small business owners and furthermore, identify how they can see success.

#### Are there any features that can help small businesses?

The provided Spending datasets show the "Change in In-Person Card Spending." This

dataset only looks at in-person card spending, and spans from 01/5/20 - 02/27/22, so the data is limited. Analyzing the spending dataset for Mattapan Square (Figure 8), we can see that this area appears to be near or at its pre-Covid % change in consumer spending. This is a healthy sign for the district; it shows that, as of February, people were willing to spend roughly just as much money in-person as they were before Covid. With most restrictions being lifted since that time, it is possible that % change levels may have even gone up since. Unfortunately, the foot traffic dataset's end date is 12/26/21 so we can not make any conclusions about how busy the area is overall at the time of the increase in % change in consumer spending. However, given the increase in spending data, we assume that foot traffic is also roughly the same as it was pre-Covid.

This data suggests that small businesses can now experience a fresh period of relief for the first time in nearly 3 years. We know that these businesses experienced their harshest times during Covid and many had no other option but to close down—after analyzing the data we can conclude that small businesses in District 4 will continue to perform better than they have in the last 2 and a half years because of the increase in in-person consumer spending. A return to higher levels of spending and traffic can draw more businesses into the district because of the opportunity that now exists.

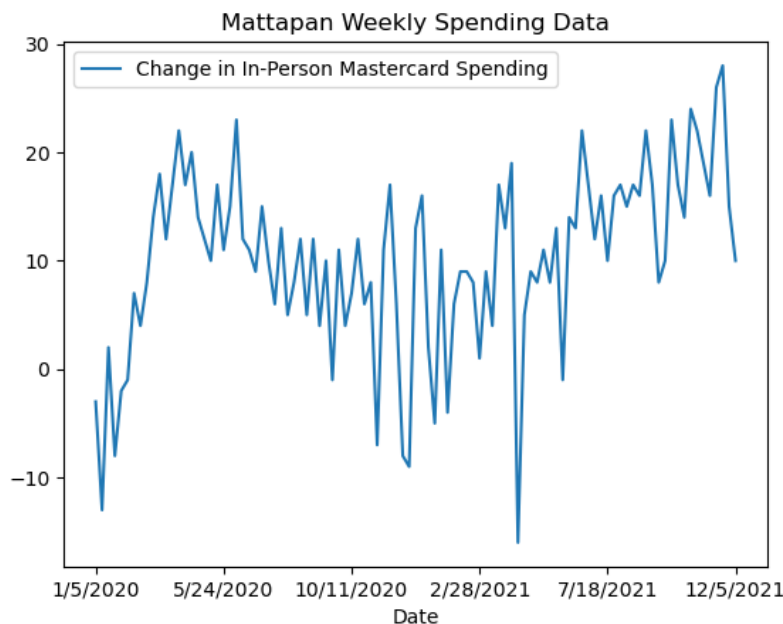


Figure 8. Mattapan Square Weekly Average Spending From 01/5/20 to 12/05/21

Jan 05,2020 until March 08,2020 [pre-covid]

March 15,2020 to November 15 2020 [during covid]

September 26, 2021 to December 05,2021 [post covid]



### Are there any barriers to a business succeeding in District 4?

Regarding the amount of capital businesses may spend, while there was no dataset with this information, we have slight indirect access to business spending through employment data from the base project (described above). From that data we can recall that District 4 had significantly lower employment than other districts at the time of data collection. From this we infer that the spending in D4 is relatively low, mostly due to a decrease in income during the low in-person consumer spending trend. Due to low incomes, a high cost of operation is not something that operators in District 4 should be concerned about. Findings on District 4's business count by category (shown above) as well as count of employment (from base project) tell us that District 4 lags slightly behind District 7, a district that has similar demographics to District 4, in both aspects. While these differences do not directly indicate that District 7 is performing significantly better, it does show that District 7 may be more appealing to business owners.

We analyzed the breakdown of employment data in neighborhoods in District 4 using the Employment Data datasets, focusing on the age category, race, education levels, etc of employees (see below for charts) to observe comparisons between District 4 and other districts.

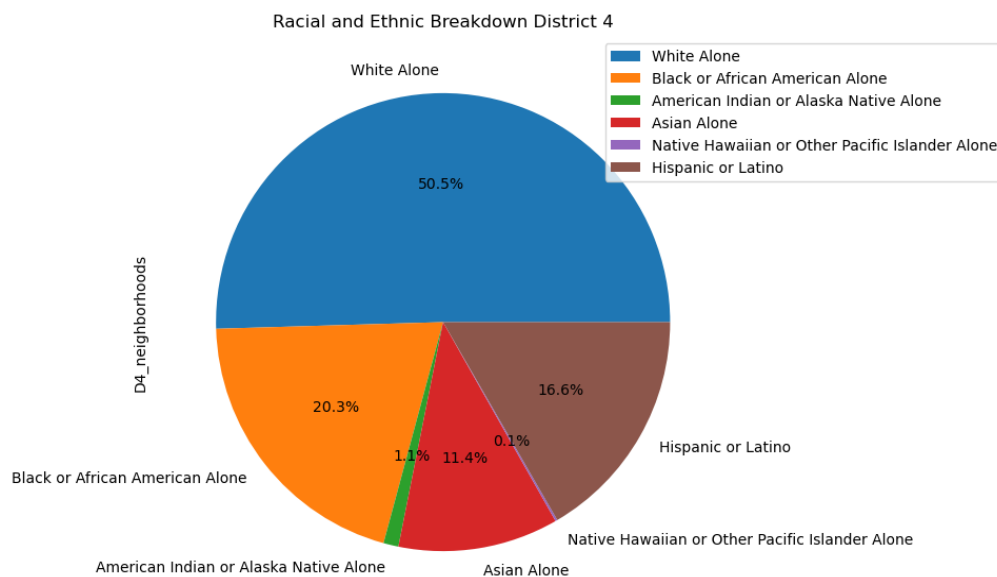


Figure 9. Racial and Ethnic Breakdown in District 4

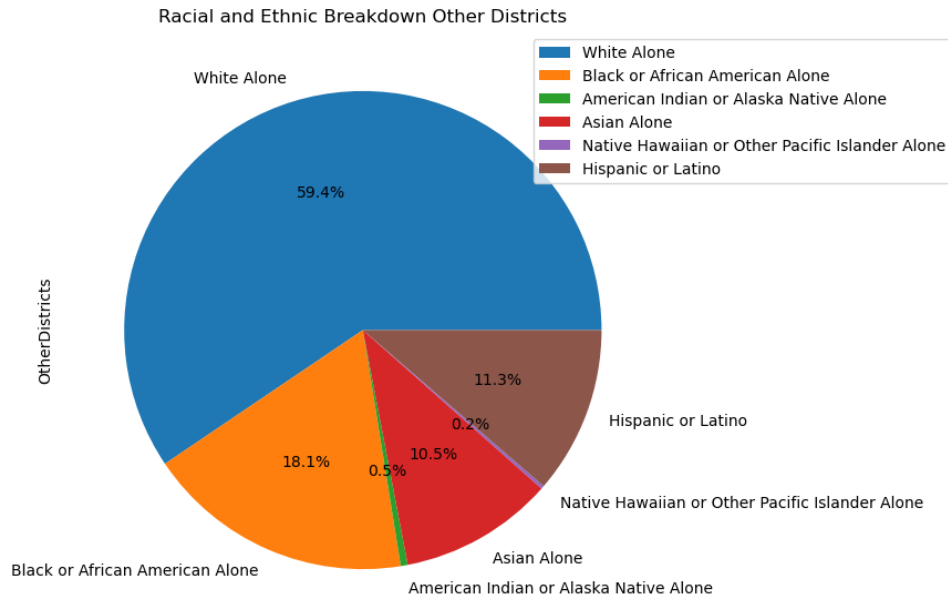


Figure 10. Racial and Ethnic Breakdown in Other Districts

District 4 has over 9% more non-White employees; this shows that District 4 is far more diverse than the average of all other districts. However, District 4 has a lower educational level overall (Figure 11). Fewer employees have an educational background higher than college with more employees having lower or unavailable educational backgrounds compared to the average of the other districts. This may be linked to the lack of educational resources in District 4 compared to the other districts as mentioned previously.

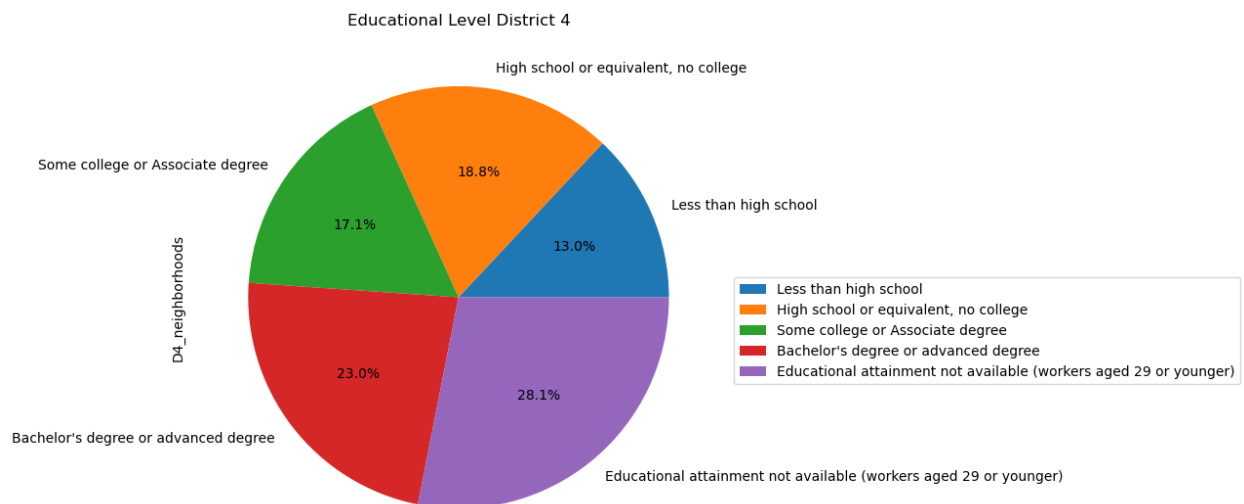


Figure 11. Educational Level in District 4

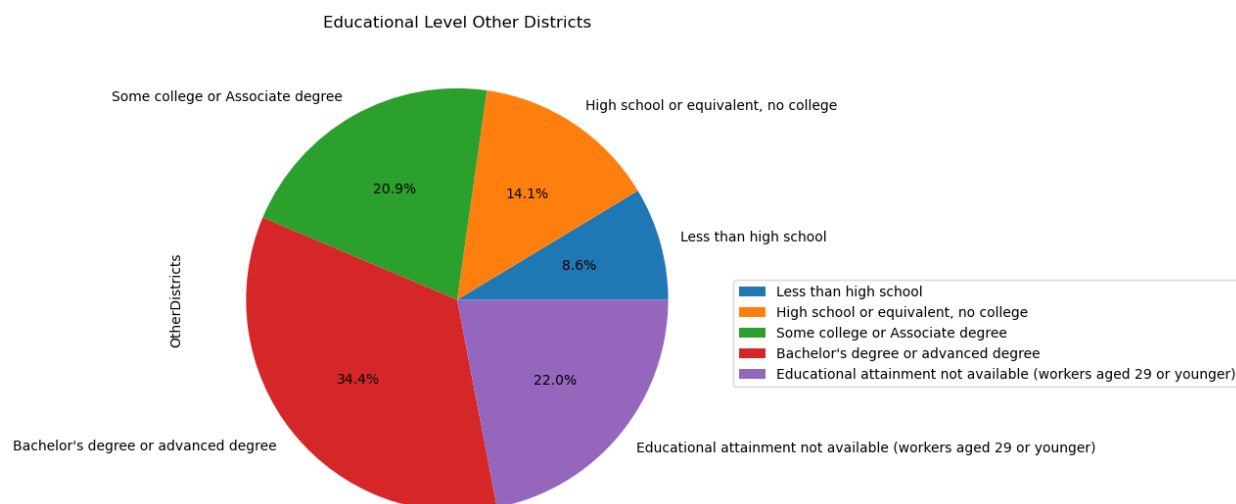


Figure 12. Educational Level in Other Districts

#### Do businesses in District 4 get visibility online?

We planned to account for visibility by finding if District 4 businesses were in datasets such as Yelp and Google Maps. To find this, we planned on merging District 4 businesses with Yelp businesses on longitudinal and latitudinal coordinates to find if any of the District 4 businesses can be found on Yelp as well as look at business ratings. Unfortunately, Yelp does not allow scraping and instead provides the [Yelp Dataset](#) for people interested in Yelp's dataset of businesses to use; however, they provide only a subset of around 150,000 businesses on Yelp. With this subset, we were only able to find 2 businesses in Massachusetts and no businesses in District 4.

If we could obtain comprehensive data from Google Maps and/or Yelp and other online services, we can observe the engagement among frequented small businesses in District 4, as well as the quality of reviews. From this, we can find if there is a strong online presence of small businesses in District 4. A high engagement would show that residents of District 4 are actively looking for the best service options in their area – they desire high quality and they want to continue to spend their money in-person in District 4. This would be something to consider for the future.

#### District 4's Business Landscape

Looking at the income of each employee, District 4 (Figure 13) has a significantly lower income than the average of the other districts (Figure 14). 37.2% of D4 employees earn \$1,250 per month or less, compared to the average of 14.8% of other districts. 24.1% of D4 employees earn more than \$3,333 per month, while 66.5% of employees of other districts earn more than \$3,333 per month. This may be related to the education level of employees and/or the business

type. The number of corporate businesses may also influence this; however, as we saw above, the number of chain stores we found per district is minimal.

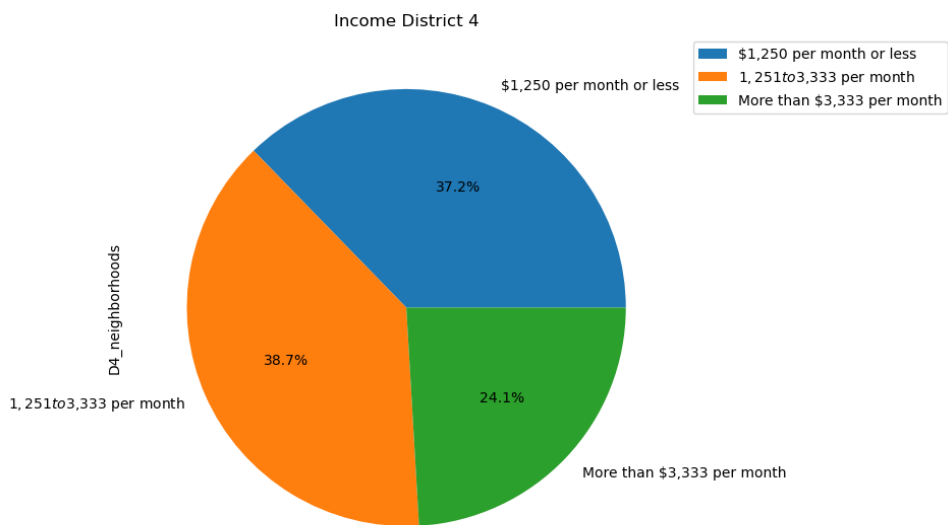


Figure 13. Income in District 4

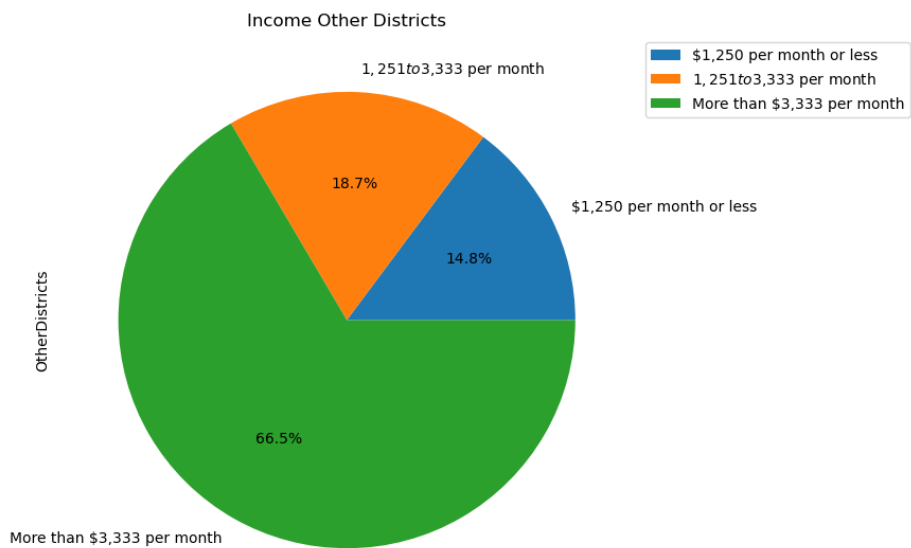


Figure 14. Income in Other Districts

It may be important to note that District 4’s customers will primarily come from only District 4 and employee income may be affected by the income of District 4 residents. District 4 may experience a lower amount of business overall because there is far less traffic than there would be in a District 2 sort of demographic (higher income, more corporate). For example, residents in District 4 or other districts are much more likely to travel to District 2 for either work or recreation than residents in District 2 are to travel to District 4 for the same reasons.

This is at no fault to District 4, but helps us understand why District 4 lags slightly behind its neighbors. District 4 is lacking in certain business types in comparison to other districts — these businesses can be focused on to further develop District 4. As it is one of the districts with a lower number of services in arts, entertainment, recreation, and education, among others, District 4 can benefit from expanding in these industries. This can keep District 4 residents within the district for these services rather than having them travel to neighboring districts like District 6 or 7, and/or attract customers from other districts to District 4.

### **Our conclusion**

From analyzing various datasets about businesses in Boston, we are able to understand the current landscape of businesses in District 4. District 4 does better in terms of race and ethnicity; however there is a lower educational level, income, and services compared to other districts. A reasonable benchmark for small businesses in District 4 may be to focus on in-demand and attractive business types, including businesses already existing in and businesses lacking in District 4, and expect neighborhood traffic and not large-scale store-front sales. With little to no barriers of entry, District 4 is favorable for small business owners looking to develop a modest yet profitable business with minimal cost of operation.

### **Challenges faced**

Challenges faced during data collection, exploration, and analysis mainly involved unsubstantial data or inconsistent information across multiple datasets. It was difficult at first to filter out which data could be useful to us and how to analyze the data, and there was limited information. Cell phone traffic data does not provide a point of reference previous to 2019 and doesn't span a substantial enough timeline or geographical area. There are also inconsistencies within a dataset – some features are not universally categorized or are categorized inaccurately between multiple datasets. Some business names are not correct in the datafiles compared to when searched online or cross-referencing multiple datasets – the data may be outdated.

### **Limitations**

It was a challenge to find Minority owned businesses after seeing such a low number of MBE certified businesses in District 4 that we filtered out from the MBE certifications dataset. The Certified Business Directory dataset is not comprehensive or reflective of current businesses in Boston. While businesses can put MWBE status on their Yelp pages, the dataset Yelp provides does not include this information and Yelp does not allow web scraping.

Some datasets are not organized by district, but by neighborhoods. Since different parts of a neighborhood can belong to different districts, the data analyzed based on neighborhoods is not accurate to a specific district. For example, foot traffic, employment data, and spending data sets are listed by neighborhood and not district. Furthermore, the datasets do not all have

a description. For example, we wanted to observe the amount of capital businesses are spending with the provided Spending Dataset. Unfortunately, the dataset does not refer to business spending but rather consumer spending at businesses, so we were unable to use this to answer the question.

### **Suggestions for the future of the project**

We are using the shapefile from before the redistricting; we believe this is important in helping understand the business environment of the original District 4. However, it would be beneficial to use the updated shape file to observe the updated landscape of small business owners in District 4 in the future. As some of the data provided were outdated or not current, more up to date data to help with finding more conclusive and relevant results. As mentioned above, with the advent of technology and increasing use of the internet, researching the online presence of small businesses and how that affects the business can be something to consider.