Team 2: Project Deliverable 1

Preliminary Analysis:

Using the currently given district shapefile, a dataset of businesses in just District 4 was found – there are 411 businesses in District 4. We performed describe() on the dataset and were given a summary of the dataset (Figure 1). An interesting thing to note from this summary is that the average number of estimated employment is approximately 8 employees, while the minimum estimated employment is 0 and the maximum number of employees is about 937.

	Unnamed:	...1	latitude	longitude	zip	NAICS_20	NAICS_20	estimated_employment
count	411	411	411	411	411	411	411	411
mean	3130.983	3131.983	42.30073	-71.0708	2122.017	571313.1	62.11436	8.216545012
std	323.9033	323.9033	0.006802	0.008512	1.120628	204668.6	17.71603	46.85774339
min	2667	2668	42.28326	-71.0848	2121	0	23	0
25%	2874.5	2875.5	42.29822	-71.0792	2121	452210	48	2
50%	2997	2998	42.29989	-71.0709	2122	541860	62	4
75%	3455.5	3456.5	42.30546	-71.0622	2122	722511	81	7
max	3697	3698	42.31362	-71.0599	2124	926150	99	937

Figure 1. district4_businesses.describe()

We looked at estimated employment per business which is the number of employees the business has the resources for. We counted the number of estimated employment per business and found that most businesses have an estimated employment of 2 - 3 (Figure 2). The bar graph shows that there is an outlier of 937 estimated_employments with the next highest estimated_employment being 91; there is a large gap between these two. This raises the question of how we define small businesses in our analysis.

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 2.1 Left Column: estimated_employment, Right Column: Number of businesses with that estimated_employment

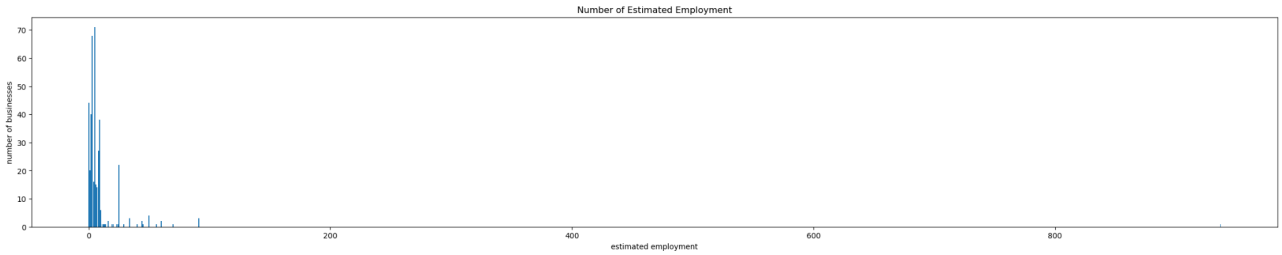

Figure 2.2 Bar graph [full size]. *There is an outlier at 937*

We found the business types in district 4 and the number of each business type (Figure 3) to answer one of the key questions.

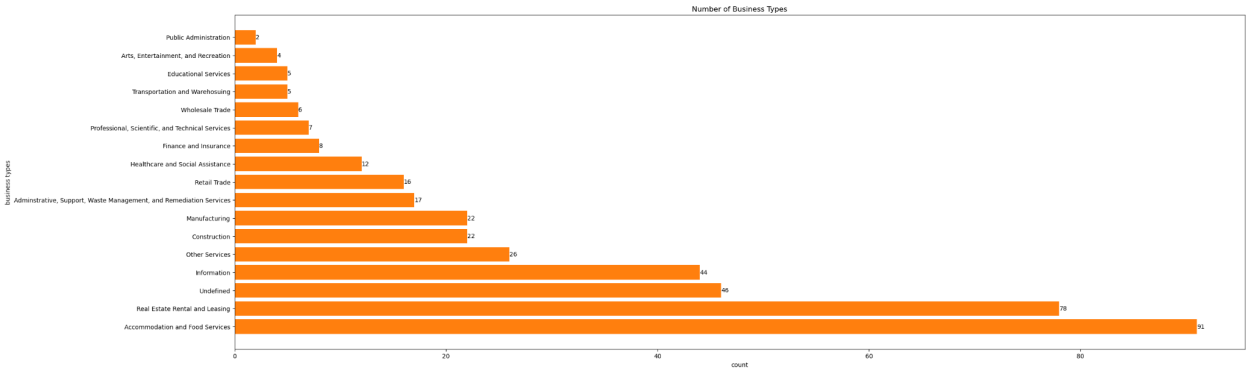

Figure 3. Number of Business Types [full size]

Key Question: What businesses exist?

District 4 has a variety of businesses including but not limited to, Full Service Restaurants, Beauty Salons, Financial Transaction Offices etc (Figure 3). The most prominent business types are food-service, retail trade, personal care services, and healthcare and social assistance. There is a lack of services in arts, entertainment, and recreation – there is one fitness center and one fitness instruction site. Notably, there is also only one gas station within the district.

With so many unique types of businesses, there are plenty of opportunities to learn from this data and be specific to each business type in our solution to the main project question. Our job with this data of businesses is to figure out how we can make it more appealing to new small business owners to enter the District 4 economy. This data also allows us to identify gaps in the types of businesses that make up the District 4 community; leading us into the right direction of which small businesses to appeal to.

Project Scope:

The scope of the project is still intact and there are no additional limitations that were not previously accounted for after seeing the data. Achieving the project goal will take further analysis of the details of District 4 itself as well as understanding the landscape that small business owners will need to navigate in order to be profitable in District 4. However, we will need to define “small business” and work on parsing the dataset with the definition. For

example, Starbucks in District 4 has an estimated employment of 12 – by the number count, this would be a small business, but we know Starbucks is a large chain, so should we be excluding Starbucks and similar businesses from the “small business” list.

Modifications to Original Proposal:

No changes have been made to the original project description.