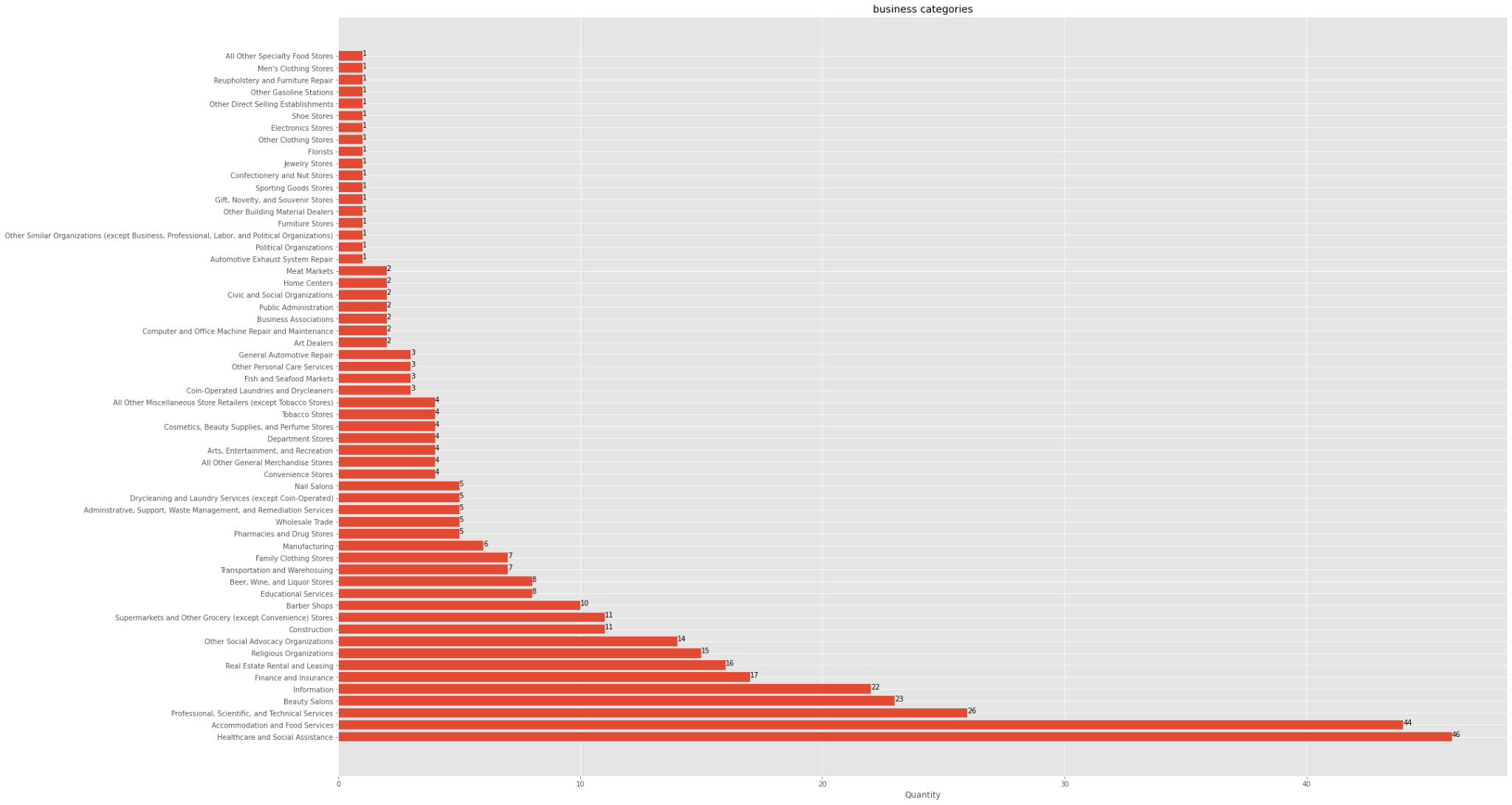
Small Business - Team 3

Galo Guerra, Jiahua Zhang, Andrew Wang, Xingru Chen

Deliverable 3

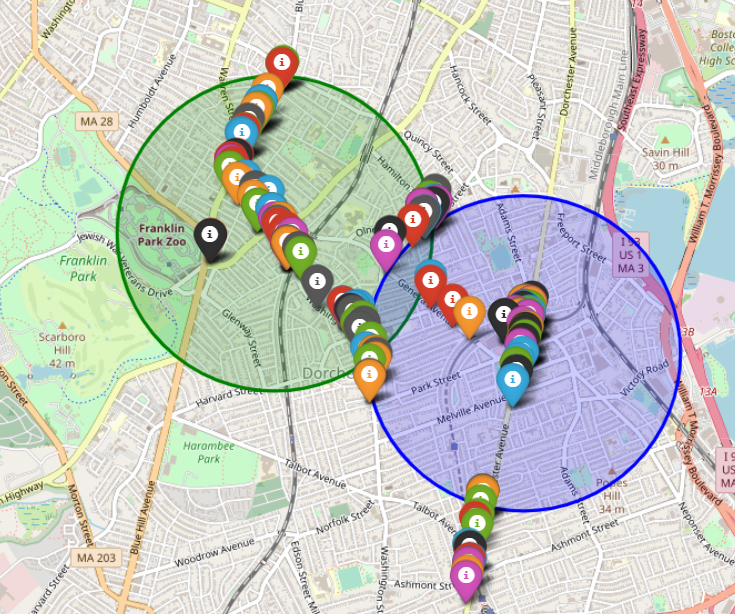
Analysis and Key Questions:

Our group was tasked with finding answers to the key questions of what businesses exist in District 4, and what businesses are overrepresented. In our research, we determined that we could categorize businesses in two ways. For most businesses, this meant using their NAICS 2-digit identifier, such as Healthcare and Social Assistance businesses. However, two specific 2-digit categories, Retail Trade and Other Services, we felt were too broad and thus split these into their NAICS 6-digit identifiers. This ended up giving us 58 different business types to analyze that were present in Boston’s District 4:

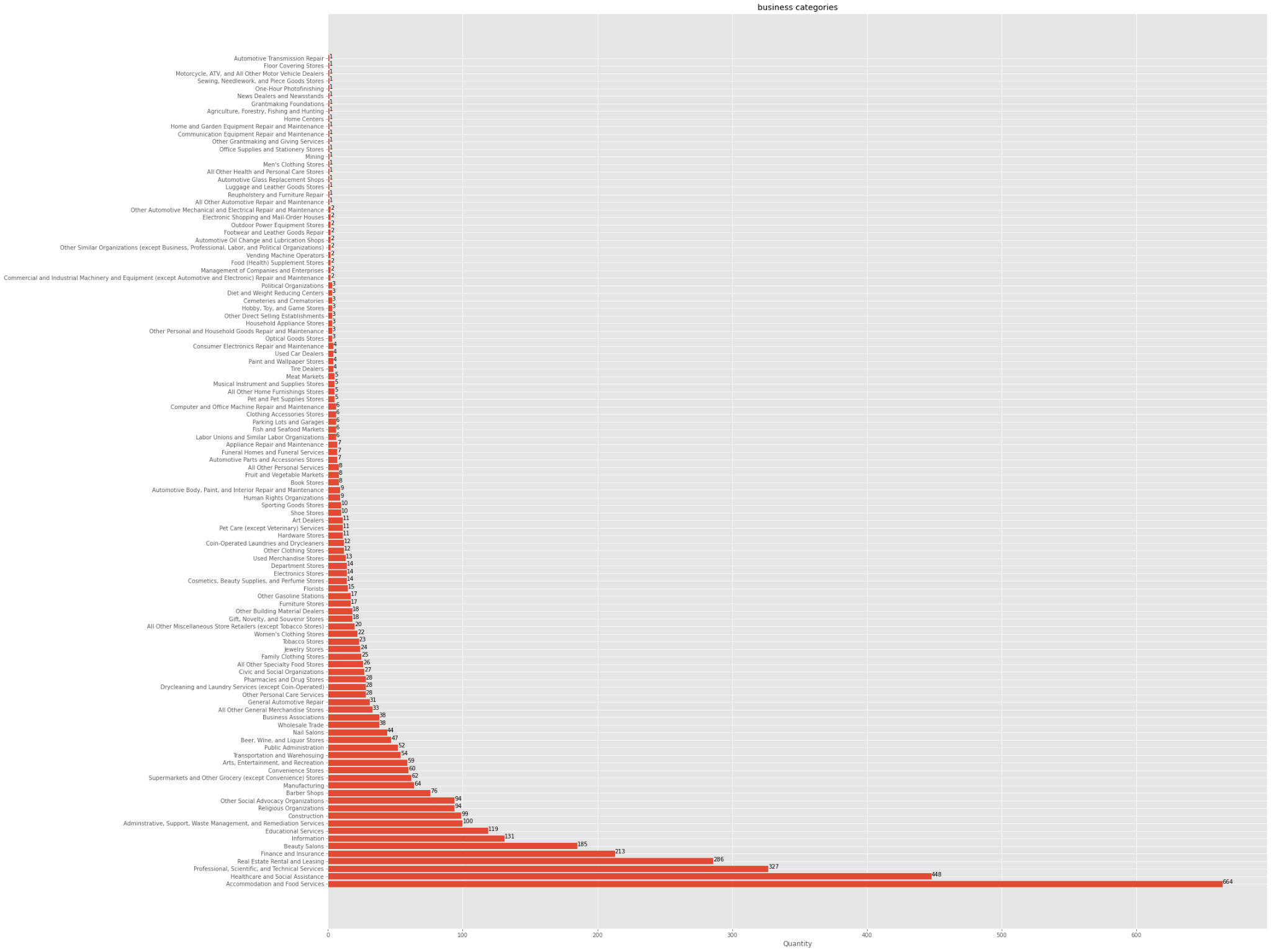


Notable amongst these businesses were the following categories: Healthcare and Social Assistance with 46 businesses, Accommodation and Food Services with 44 businesses, Professional, Scientific, and Technical Services with 26 businesses, Beauty Salons with 23 businesses, and Information with 22 businesses. These businesses are the most represented within District 4, and are all contenders for being overrepresented in District 4. However, most notable is Beauty Salons, as this is the only 6-digit identifier that is in the top 5. This is probably the most clear answer for what business is overrepresented, as it is the only very specific identifier that contends with the less specific identifiers numerically.

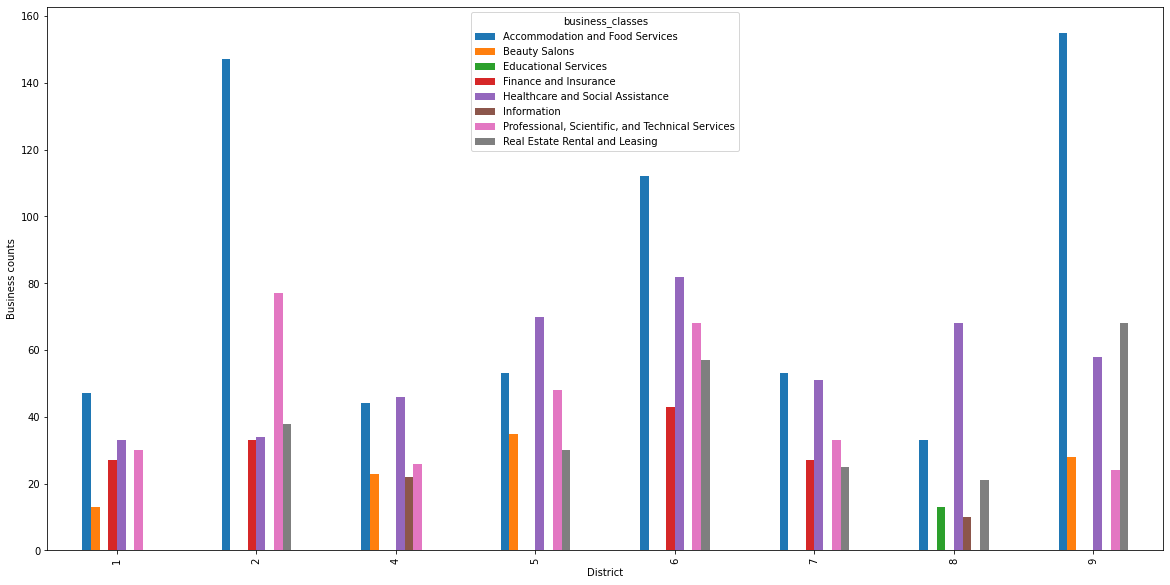
From there we mapped the most common businesses within district 4, assigning each to a color to provide a clear view of how these businesses are laid out. As well, we created two centroids in order to provide a clearer picture of how close the business types are to one another:



From there, we were able to compare the results in district 4 to those in other districts. In comparison to the other 8 districts, we see a lot of similarities in the most frequently seen businesses. We created a graph to see the number of businesses in all other districts first:



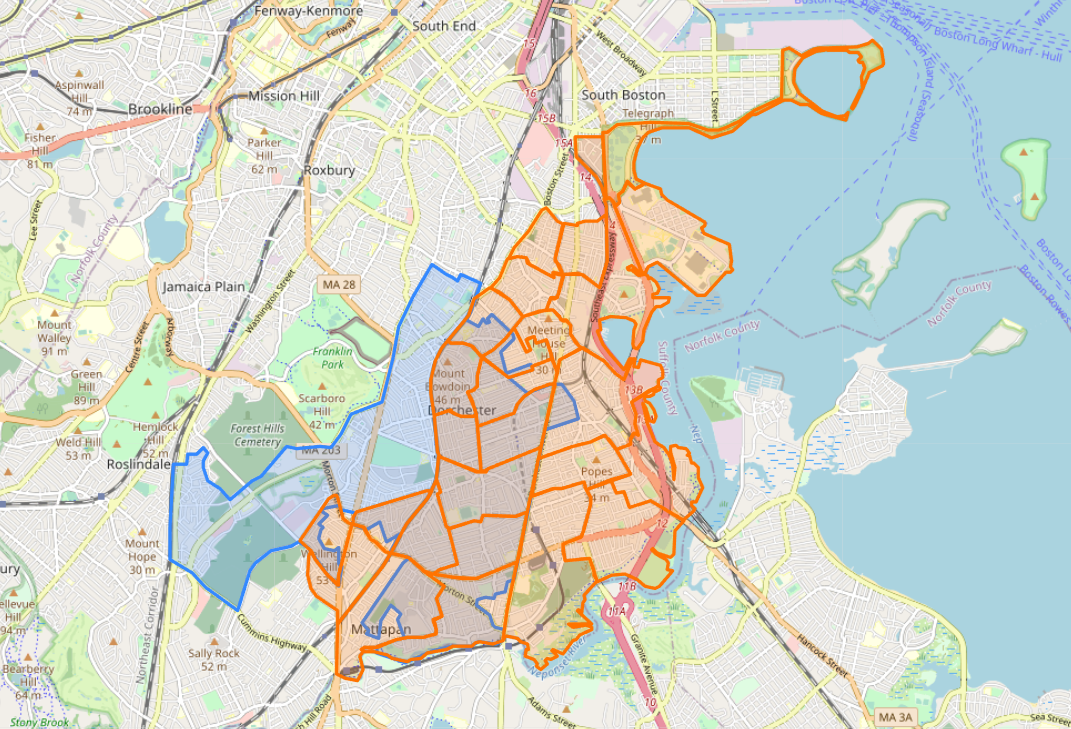
Fram there, we separated the largest categories and created a multiple bar graph that shows every district, and the five most represented business types in every district:



Notably, all districts share: Accomodation and Food Services, Healthcare and Social Assistance, and Professional, Scientific, and Technical Services. As well, districts 1, 5, and 9 interestingly share the category of Beauty Salons. This was surprising as it appears that maybe Beauty Salons could be viewed as overrepresented over multiple districts.

Extension Project:

For our extension project, our group chose to look at Boston’s 2010 Climate Social Vulnerability Data. This dataset provides information on different types of vulnerable populations in Boston (such as the elderly, children, minorities, low income households, etc.), and provides their locations based on the 2010 census tracts. We were able to use this data to map out the census tracts located in Dorchester and Mattipan, the two largest parts of District 4:



As you can see by the orange sections, much of the east part of district 4 contains at least some form of vulnerable population. Using the dataset, we were also able to find that within Dorchester and Mattipan, there is the following vulnerable populations breakdown:

|  | Children | Elderly | POC | Limited English | Low to No Income | People with Disabilities | Medical Illnesses |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Dorchester | 17,424 | 6,535 | 50,778 | 29,284 | 22,749 | 9,819 | 25,010.76 |
| Mattipan | 8,859 | 4,073 | 28,276 | 13,604 | 9,531 | 5,605 | 12,396.87 |

Our current focus is on continuing our work in the extension project, in two separate ways. The first, and more obvious way is by introducing more restrictive use of the dataset, as we are currently viewing all of Dorchester and Mattipan, while only some of the neighborhoods are contained in District 4. So we hope to refine our map, and get more accurate numbers as to the number of vulnerable people living within District 4. The second way we intend to improve our extension project is by introducing our original key questions into it. We want to find how many businesses exist within the highlighted orange areas of district 4, and is there any insightful data we can find from the types of businesses in these areas. As well, are any businesses too prevalent or completely non-existent in vulnerable populations?