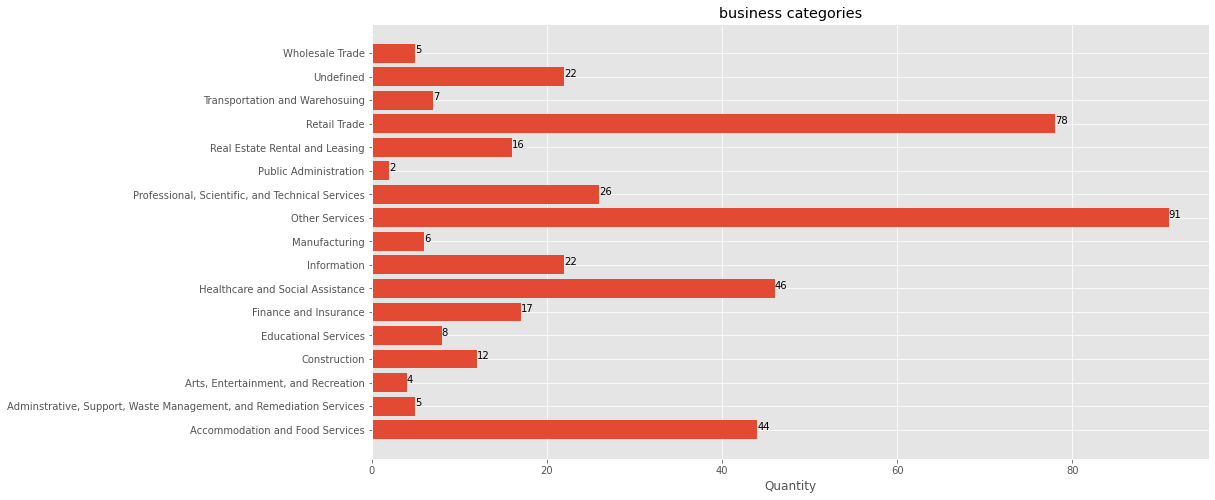
Team 3

Preliminary Analysis:

Our preliminary analysis of the data has indicated that there are 411 applicable businesses in District 4. Each row contains address information, type of business, NAICS information, and employment details. Our group first used geopandas to limit the information to district 4, before using folium to apply it to the given shape file. From there we were able to create a chart of the different business types:



We can see that there are a wide variety of businesses, with some types being represented as few as 5 or less, and some of over 40 businesses. So far, this has given us a basic idea of the number and types of businesses we will be analyzing to fulfill this project.

Key Questions:

The key question we have tackled is what businesses exist and which are missing. We used the graph above to answer these questions. Though, what may not be realized at first is that the largest individual category at 91 businesses is actually “Other Services” which includes a variety of business types and we will need to break down going forward. However, we can notice that Retail, Healthcare, and Food Services are in far greater numbers than other business types. Going forward it seems we should be looking to expand the underrepresented business types such as arts and entertainment, public administration, manufacturing, and wholesale trade.

Project Scope:

From our current analysis, we have found that a vast number of the businesses are undefined or labeled as other services for their category. This limits our categorization and we would potentially want to further redefine these vague categories. There are still some pieces of information about businesses that we would like to obtain, such as ownership. From an initial glance of the data there are some businesses that may not be classified as small businesses such as Xfinity and 7-Eleven. We would need to determine whether we keep businesses of this type.