Background

The last few years have shaken up how businesses operate. Small, independent businesses are most vulnerable to instability and can be overshadowed by the wants and needs of their much bigger competitors. In this project, we will investigate the current situation of small businesses in District 4 and determine the factors that attract them to the area. We hope by doing this project, we can strength and build opportunities for communities in District 4. Also, we hope to provide helpful insights on the challenges that small businesses face, where they are succeeding, and how we can promote more of that success.

To complete this project, first we need to understand the current landscape of District 4, which means we need to identify and analyze the current gap and the current issues small businesses face. Next, we need to determine what will attract small businesses to the area. To work on the project, we meet twice or three times each week, depending on our progress. The meet somethings is in person, sometimes via Zoom. We carefully looked at our work together each time our group met and discussed it. Then we divided different jobs based on personal interest, and the group leader coordinated and assigned tasks to ensure the workload was fair and efficient.

Because there is no clear measurable definition of some key concepts, we give some quantitative definition of some of them according to some research. We define small businesses as "companies that have less than 25 employees." We also defined overrepresented as industries with a 25% increase in the business count by industries compared to other districts. We decided on less than 25 employees for the small business definition because that limit captured most businesses within the data and also excluded huge companies with multiple branches, after skimming through the data. While we were worried that chains and franchises would find their

way into this definition, we decided we were okay with that due to the fact that after thorough online research, we realized most of these chains and franchises were operated and ran like small businesses, but given our data limitations, it would have been an arduous task to research every business that we analyzed. Over 90% of the businesses we considered had under 25 employees. As for the definition of over-representation, we were severely limited by the data and simply chose it because it was enough of a separator value when we compared to other districts and looked at Boston overall.

This report contains two major parts: the base project and the extension project. The base project is where we analyze the given district shape file and Boston Planning and Development Agency (BPDA) business listings to understand the current situation small businesses in District 4 face. We would focus on answering these two questions: "What businesses are over-represented in District 4" and "How are businesses doing based on foot traffic." On the other hand, the extension project aspect of the report is where we expand on the analysis of small businesses based on our self-defined topic. We would look into the impact of Covid-19 on small businesses and the methods they used to face Covid-19.

What businesses are over-represented

There was a fair amount of work in terms of exploring the data and defining terms and building accurate interpretations when answering the question: "What businesses are over-represented in District 4". Some of the first problems we encountered was an appropriate definition of small business? Although the federal government interpreted small business as a business with less than 500 employees, this definition did not help us come to a respectable conclusion or solution. We defined a small business as a business that employed less than 25

employees. Even with this definition, we are unsure if it helps us truly understand the landscape of businesses in District 4. We do not believe that employee count should be the only factor involved, however, given the limitations of the data, this was the best factor we had. Once we processed our data to only include businesses that fit the small business definition, we had to determine which district each business was within. Once we were able to separate the businesses into their respective districts, it was time to see how District 4 fared. We grappled extensively with how we should compare District 4 to other districts. Although each district had a relatively similar population, they varied drastically in business count, which would surely impact our interpretations. Moreover, was it fair to expect districts to have similar amounts of different business types? We decided to limit our exploration to the districts with similar business counts. We compared District 4 to Districts 3, 5, and 7.
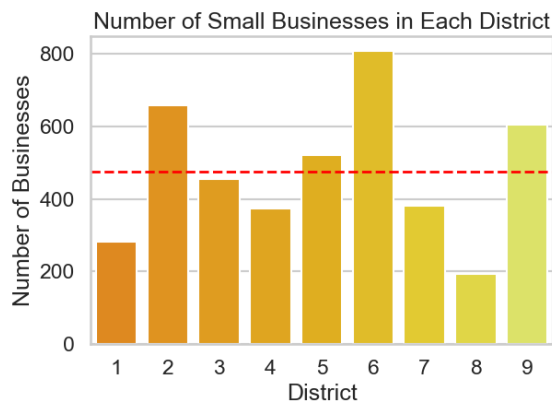

We found that the most prevalent businesses in District 4 were beauty salons, full-service restaurants, religious organizations, other social advocacy organizations, and wireless telecommunications carriers (excluding satellite). Compared to Districts 3, 5, and 7, District 4 had the lowest representation of full service restaurants, the highest representation of wireless telecommunications carriers (excluding satellite), it had approximately the same amount of religious organizations, which means it had a higher representation of them, due to the fact that

there were in general less businesses in District 4 compared to districts 3, 5, and 7. Finally, the

representation of beauty salons and other social advocacy organizations is in line with the other

districts so they are fairly represented.

Count of Businesses by Business Type and District
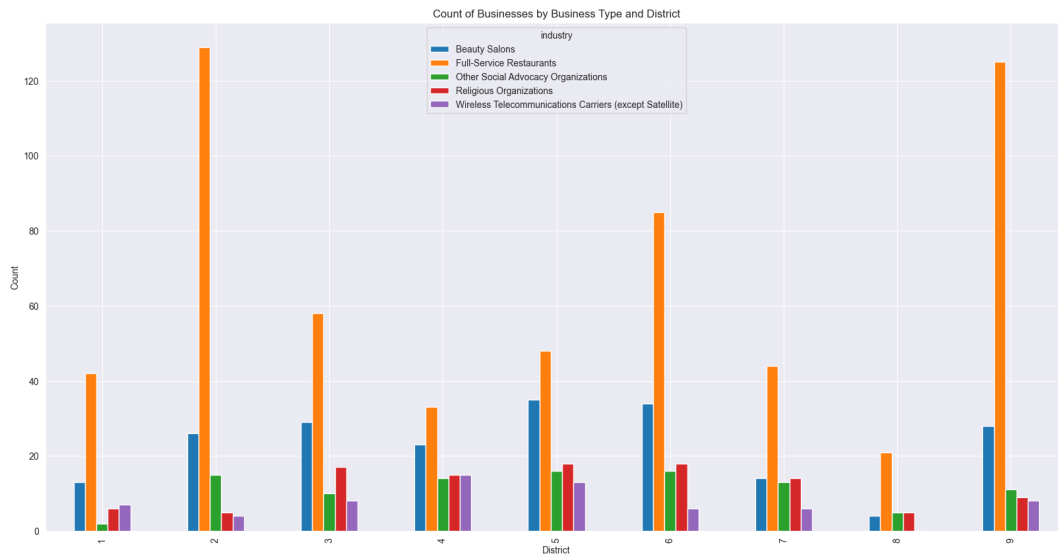
We offer a visual of the 25 most common business types within District 4:

Business Count by Business Type (TOP 25)
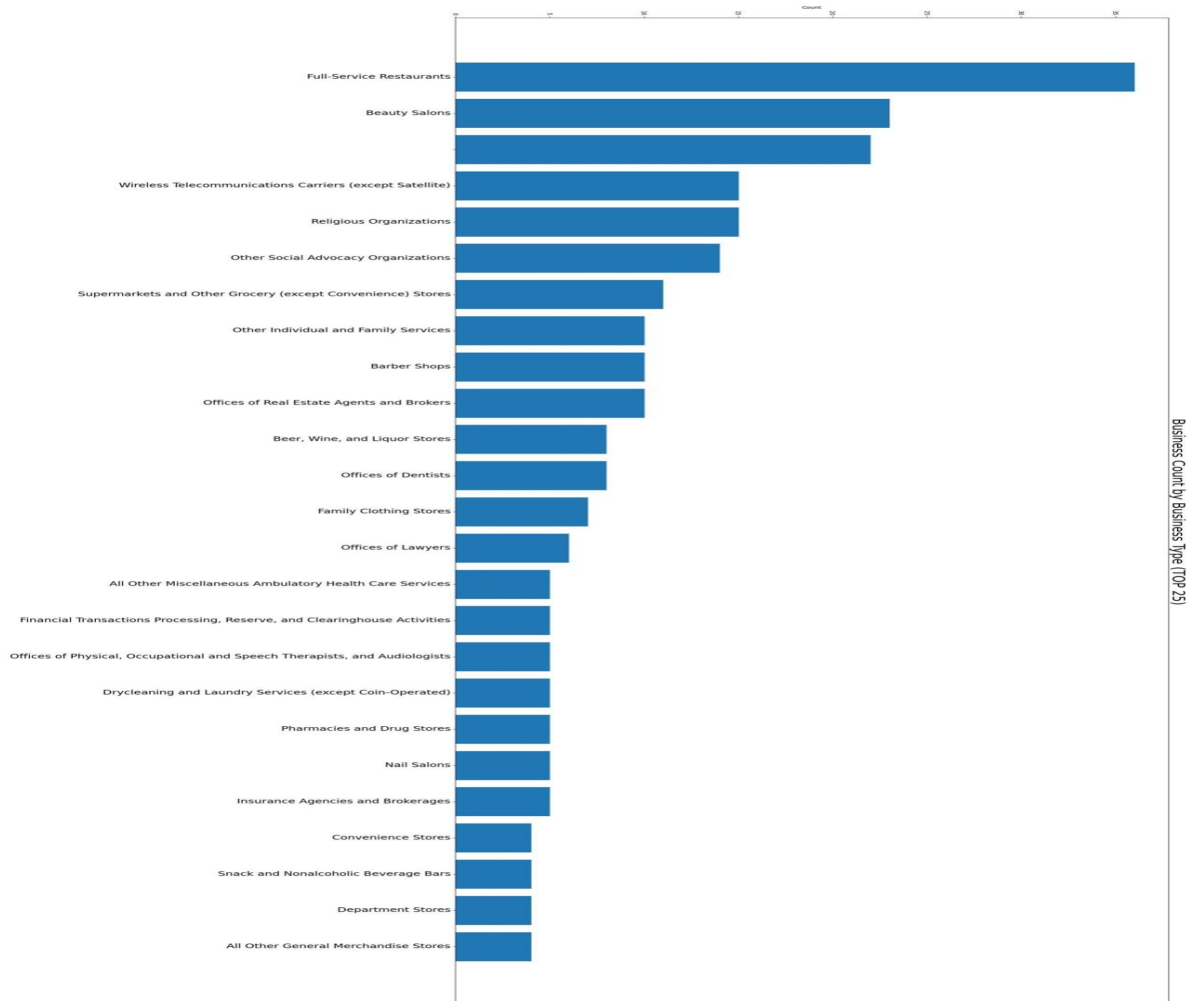
If we look at business representation from the perspective of how close businesses of the same type are in comparison to each other, within District 4 we can see that businesses tend to exist within .75 miles of another business of the same type. Since this is the range of most businesses, we will say that business types outside of this range are either over-represented if they are closer, or under-represented if they are farther. This definition may be too much of a blanket statement and with more data, we may be able to reach better conclusions with better definitions. Moving forward with this definition, it is best to interpret this definition through the lens of a consumer of these business types. For example, commercial banking businesses are on

average 1.25 miles away from each other. This may mean that there is not a suitable market for more commercial banking businesses within District 4.

Average Distance Between Businesses by Business Type
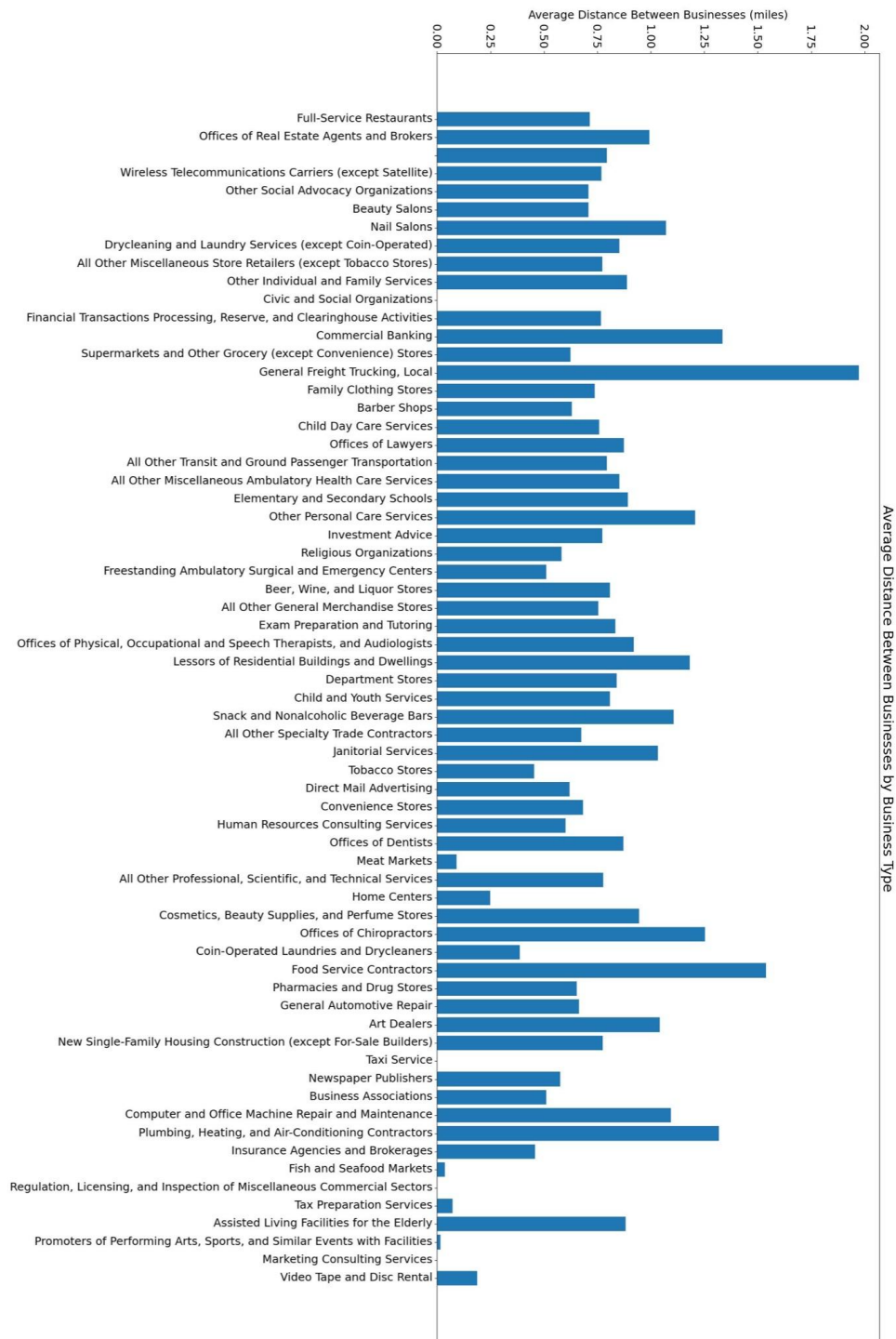
In conclusion, a few insights we can offer through this definition are the following: under-represented business types in District 4 can be categorized into three groups: niche businesses, retail businesses, and commercial businesses. Niche businesses, like offices of notaries, refrigerated warehouses, and software publishers. These businesses are not that common because while they are important, they are not as frequently utilized and so the market of supply remains low. The retail businesses is a shocking category considering how residential District 4 is. We would have expected that jewelry stores, electronics stores, and shoe stores have higher representation in District 4 but it is not the case. Maybe day-to-day consumers are shopping in Districts with more overall businesses due to the convenience of finding everything you need in one central location. Finally, the commercial businesses category is extremely important to helping us understand the problems with businesses in District 4. There is an under-representation of businesses like commercial banking, janitorial services, and food service contractors. This is a problem because commercial businesses that serve the general public rely on these services—if businesses have to go through more trouble to find businesses that can help make them successful, they will be less encouraged to operate within District 4. This is an even bigger problem for small businesses that typically have smaller management teams to operate every aspect of the business. Unfortunately, this is a destructive loop: if businesses aimed at other businesses do not exist or are harder to come by, less businesses will choose to operate in District 4, de-incentivizing the businesses aimed at other businesses from operating within District 4 as well.

Foot Traffic Analysis

The data of foot traffic comes from BPDA dataset, which each file presents the percentage change of the foot traffic compared to last week in each main street in the Boston area. To analyze the foot traffic in District 4, the file corresponds to the mainstreets in district 4 selected and combined, and using the mean of all mainstreets to represent the foot traffic in district 4 as a whole.

For visualization, the following bar chart shows the monthly percentage change of the foot traffic in district 4. In this graph, it is observed that there is an enormous decrease in foot traffic, and the foot traffic has been negative for more than two years, which indicates that the foot traffic has kept decreasing.
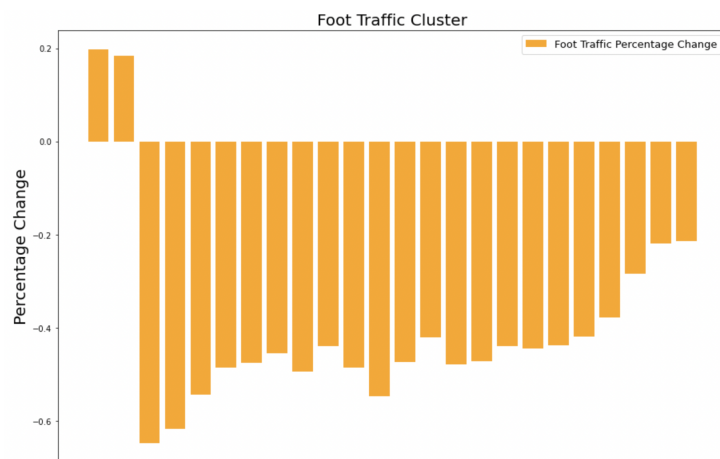

To further identify the trend of foot traffic and what happened to the small businesses, elbow method is used and it is suggested that there are mainly three big clusters of foot traffic, and the data is clustered using KMeans clustering. The clustering result is visualized as follows. This clustering
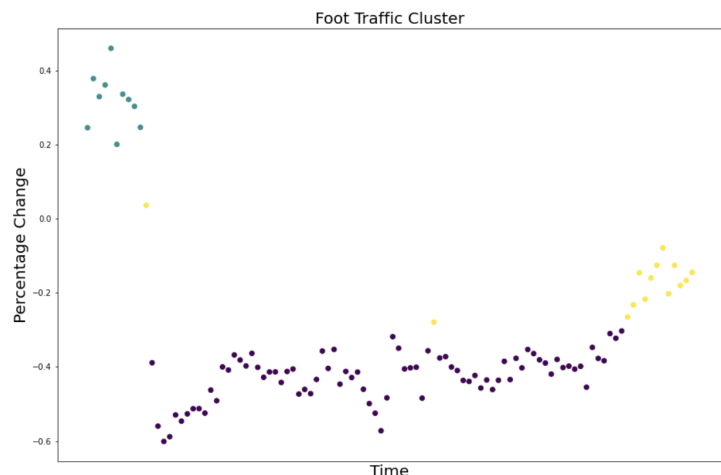
result points out three stages of foot traffic from 2020 till now, which can be explained by covid-19, and this result also gives a view of how small businesses are doing during covid-19.

In the graph, three stages are identified. The first stage is before March 2020, which is the stage before Covid-19, a positive trend is shown in this stage as the percentage change of foot traffic is always positive, which indicates the increase in foot traffic. However, after March 2020, the second stage comes, which corresponds to the beginning of Covid-19. At this stage, as shown on the graph, there is an immediate and enormous decrease in foot traffic, and this shows the huge negative impacts Covid-19 has brought to small businesses. During this stage, there is slight change in the percentage change on foot traffic, and the foot traffic remains negative, which shows the foot traffic keeps decreasing. After August 2021, the third stage comes, which is the recovery stage. In this stage, the foot traffic keeps decreasing, but at a slower rate as the percentage change is approaching zero. However, it is still far away from the foot traffic before covid, and the speed of recovery is relatively slow. The foot traffic data indicates the huge impact from covid to the foot traffic, however, the decreasing foot traffic does not completely mean the recovery progress is far behind, because during covid-19, some businesses might not able to survive due to the pandemic, while some of them might changes their way of providing services, for example, they might go virtual or hybrid, as they might provide their services through online platforms.

After analyzing the foot traffic data of the whole boston area as shown on the graph, it is shown that not only district 4 but also the other districts in Boston is under negative impact of
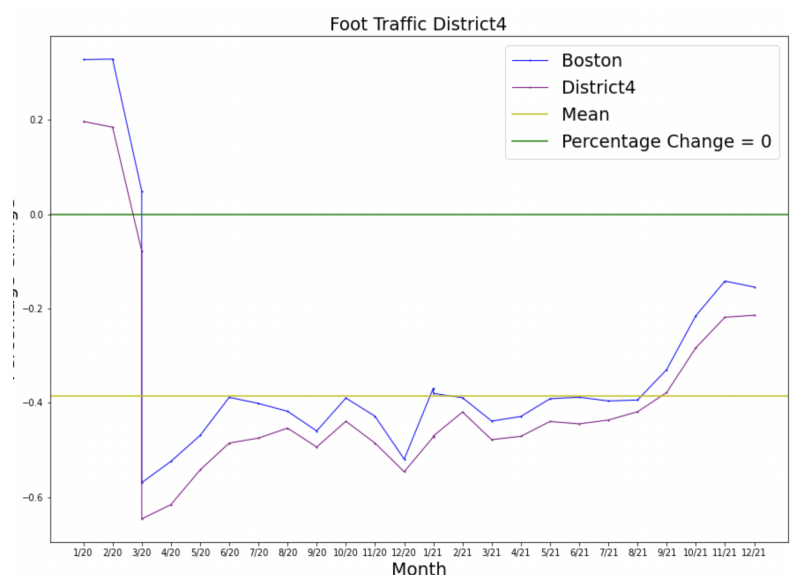
COVID-19, and district 4 is following the same trend with Boston Area. However, the foot traffic of district 4 is 14.37% behind the average of the whole Boston area, so the foot traffic of small businesses of District 4 can also be improved through other aspects.

Limitations

Lack of data is the major limitation. For example, for the data on foot traffic, the data starts only after January 2020, and is not updated to the last. Therefore, we are not able to look at the foot traffic before covid, and not able to analyze focusing on other aspects other than Covid-19. Since there is no long-term data about the foot traffic before Covid-19, we assume Covid-19 has a major responsibility for the change in foot traffic.

Also, there is no clear definition of some key concepts like small businesses and overrepresentation, so we could only give our own definition, and this might lead to some confusion and our result might include a slightly personal perspective because this is the definition we choose.

Extension Project

Through exploring the base project, we believe Covid-19 plays a huge role in the negative foot traffic of small businesses in District 4. Therefore, we want to know how did small businesses adapt, whether it is going virtual by having delivery options, outsourcing to delivery platforms, or changing their business types or products sold. We also want to know how much support did the government provide to small businesses. By knowing the answers to these questions, we can understand the kinds of government support that are useful, which may be a way to attract new businesses to District 4.

First, in order to have an understanding of the monetary support provided by the government, we collected data on all Covid-19 government spending of small businesses in District 4 from USASpending. After analyzing the data, all of the Covid-19 government spending on small businesses comes mostly from the Small Businesses Administration (SBA), and compared to SBA other agencies only spend small amounts to support small businesses. SBA has spent $2,210,005,126 in total on Covid-19 spending, and the other agencies spent $730,906 in total. Therefore, the support of small businesses for Covid-19 mainly comes from SBA.
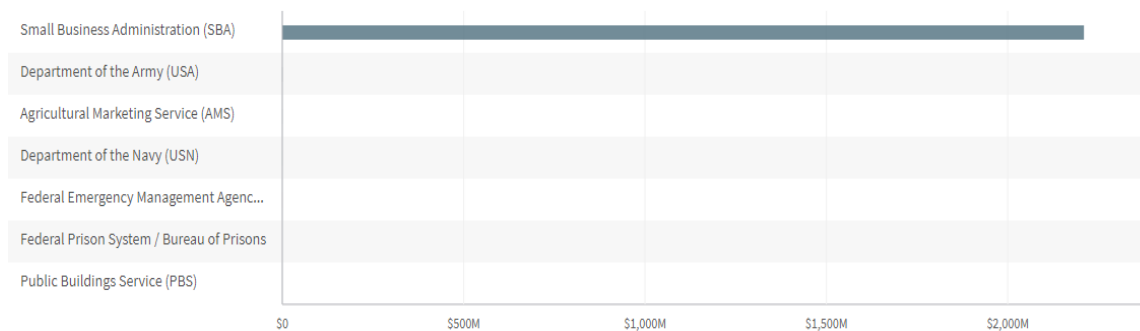
From the dataset, we also observe that among all the industries, the research and development industries get the most support from the government, and it is a lot more than others. The chemicals industry and the household appliances industry are the top two and top three industries that receive the most support. The other industries receive less support from the government for Covid-19.
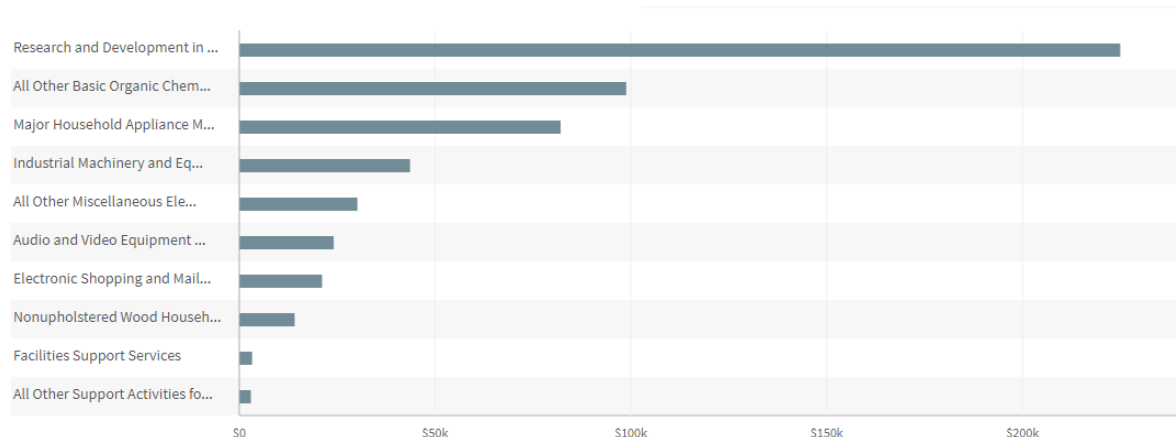
Moreover, the following is the visualization of the amount of government spending on Covid-19 small businesses. From the graph, we can see that the quarter 3 of 2020, which is about three months after Covid-19 starts, has the highest amount of government support on Covid-19. The quarter 2 of 2021 is another peak of government support, this might be because the government spending is offered to small businesses by year. After 2020, the amount of government support on covid to small businesses has decreased year by year, and there is almost no government support on Covid-19 right now. Connecting to the earlier analyze on footprint, this might be a reason why the recovery process is slow as there might be considered as no need to provide financial support to small businesses has Covid-19 is considered as finished, however, the small businesses are still at the stage of recovery, and they are still under the impact of Covid-19.