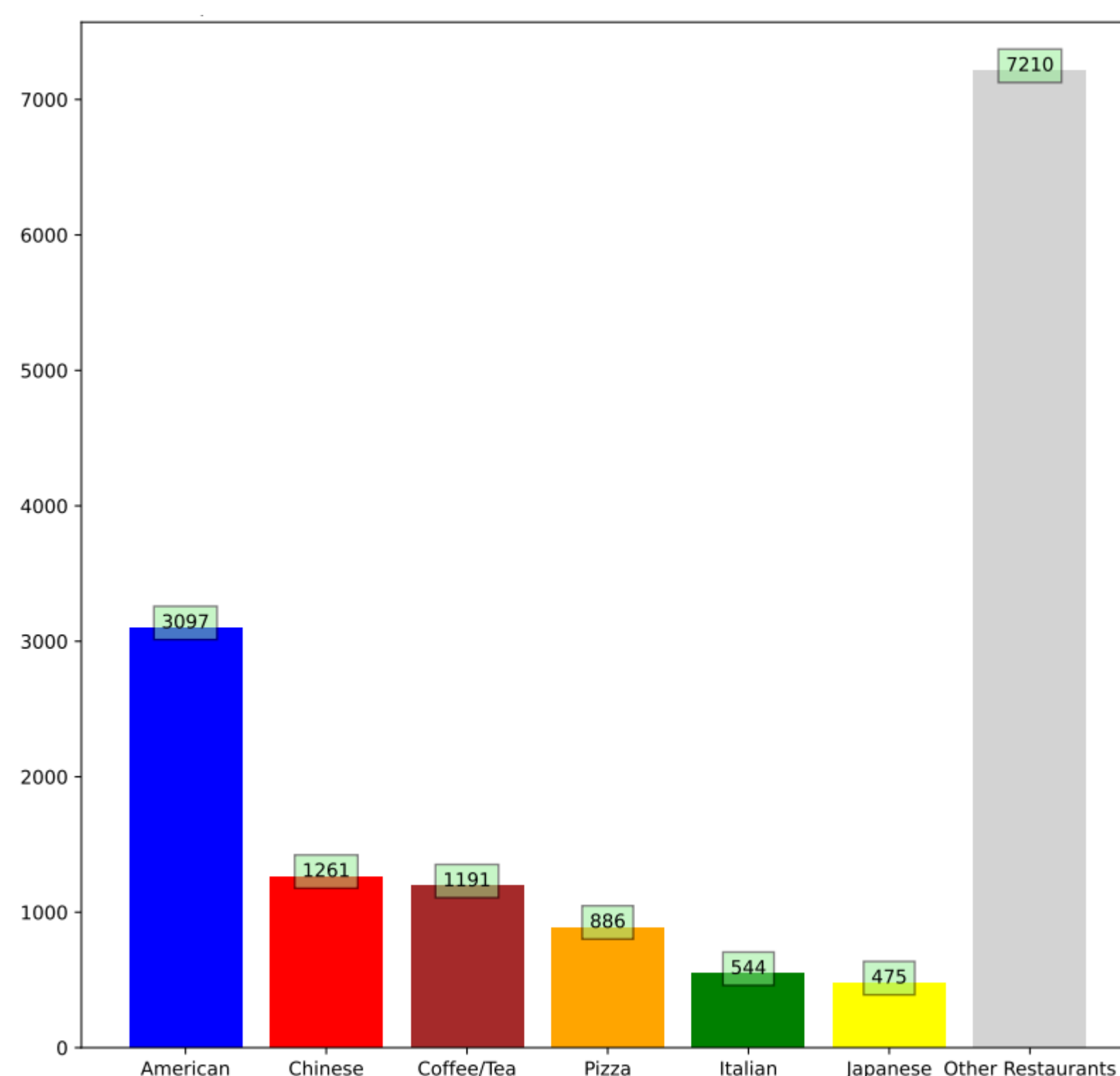


About the Dataset

The dataset we are using to perform data analysis is the New York City Restaurant Inspection Results from the New York City Department of Health and Mental Hygiene (DOHMH). The dataset contains every sustained or not yet adjudicated violation citation from every full or special program inspection conducted up to six years before the most recent inspection for restaurants and college cafeterias in an active status.

The dataset contains 331372 rows and 26 columns. Each row represents an inspection record on a restaurant and its information. Hence, one restaurant might have many inspection records. For the columns, we are interested in restaurants types, addresses, geographic locations, and violation related information.

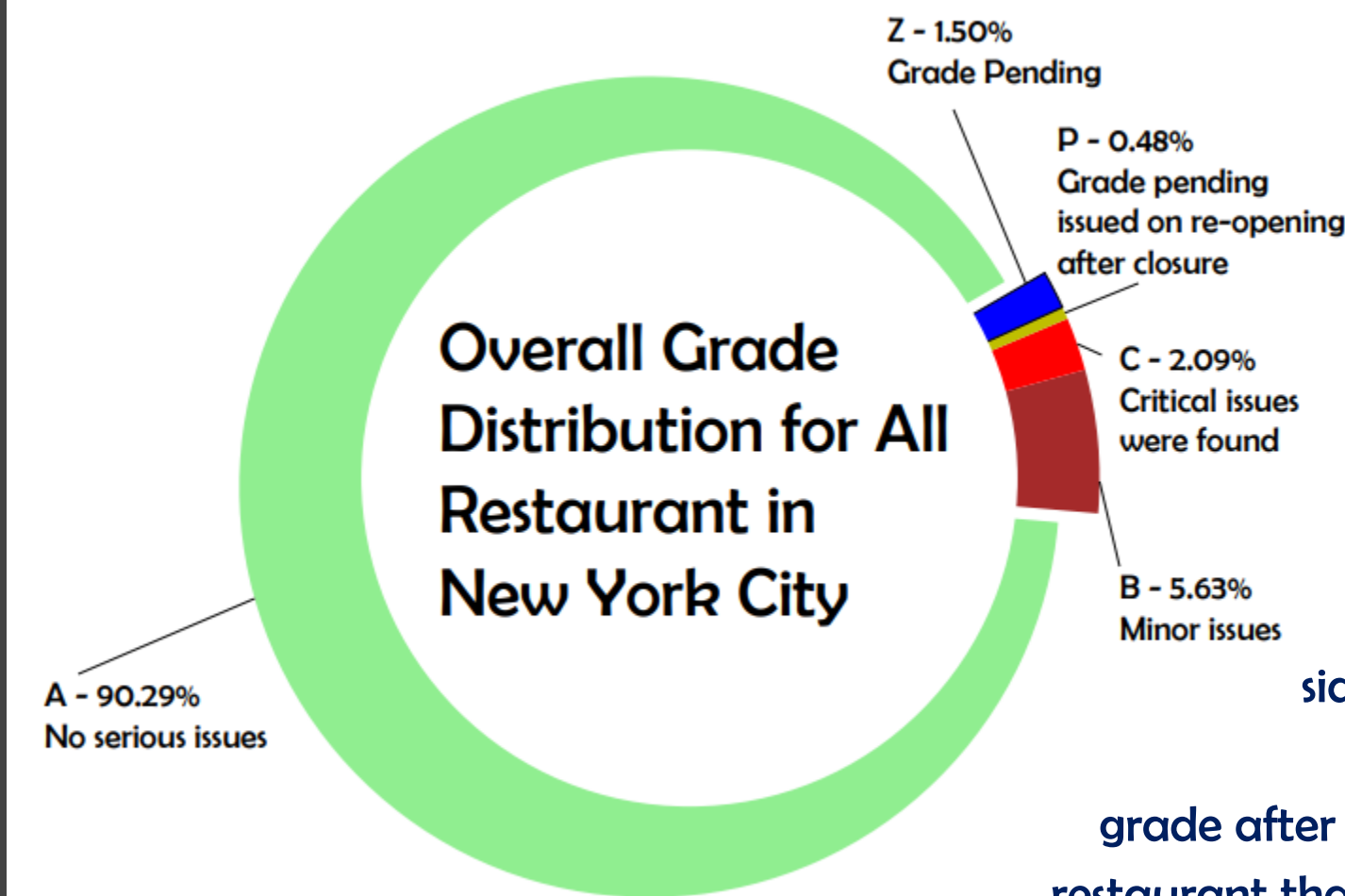
Top Six Most Common Restaurants and The Total Amount of Other Restaurants



There were total 84 types restaurant in the inspection records, and this bar plot gives us an overview of the restaurant types in New York City.

American food restaurant is the most common restaurant we can find in New York City according to this plot, followed by Chinese food, Coffee/Tea, Pizza, Italian food, and Japanese food restaurants. The top 6 restaurants added together is as many as the sum of the rest 78 types of restaurant.

The Grade Distribution Summary



Restaurant in New York City

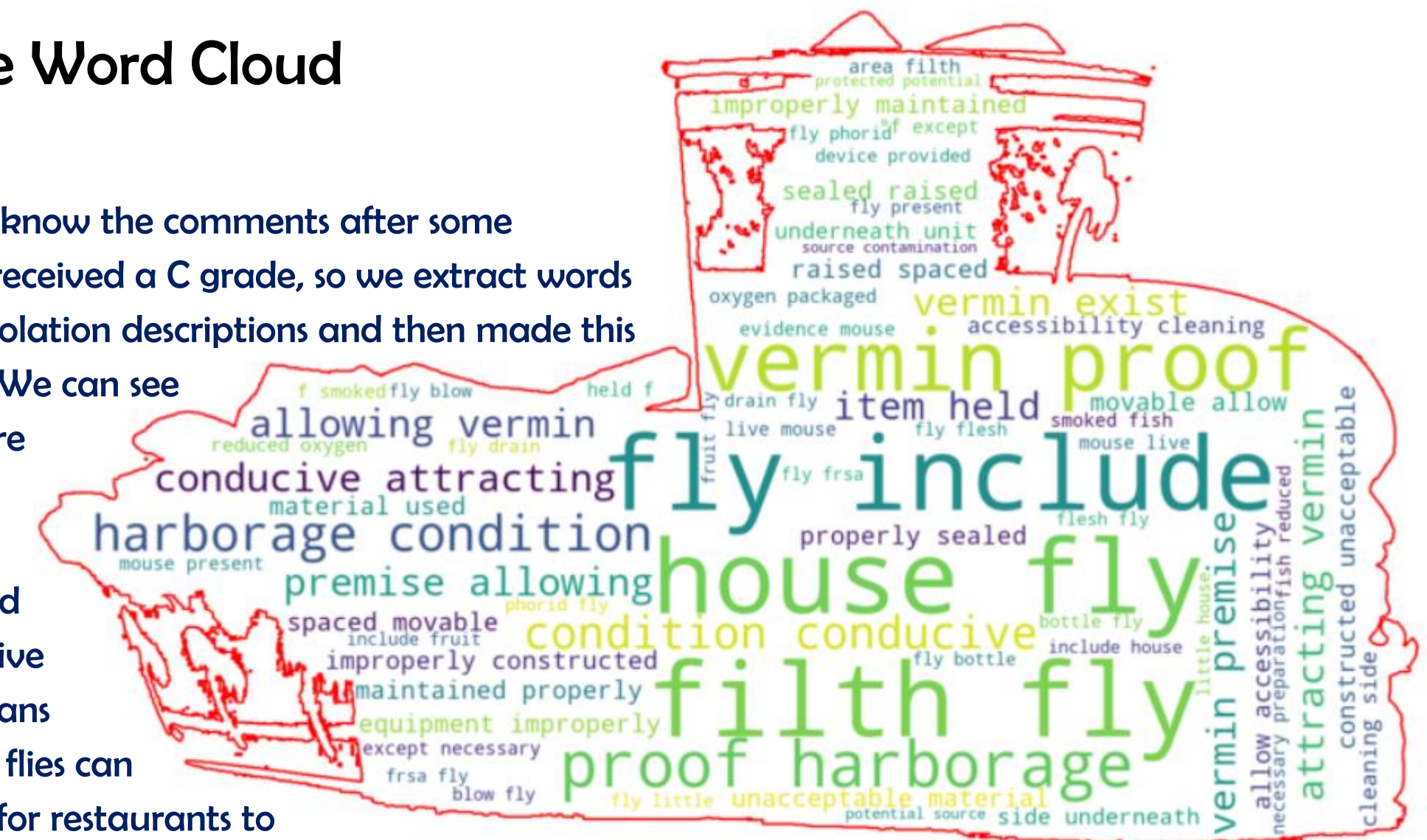
Grade	Percentage
A	77.28%
B	5.63%
C	17.09%

serious issues

B - 5.63%
Minor issues

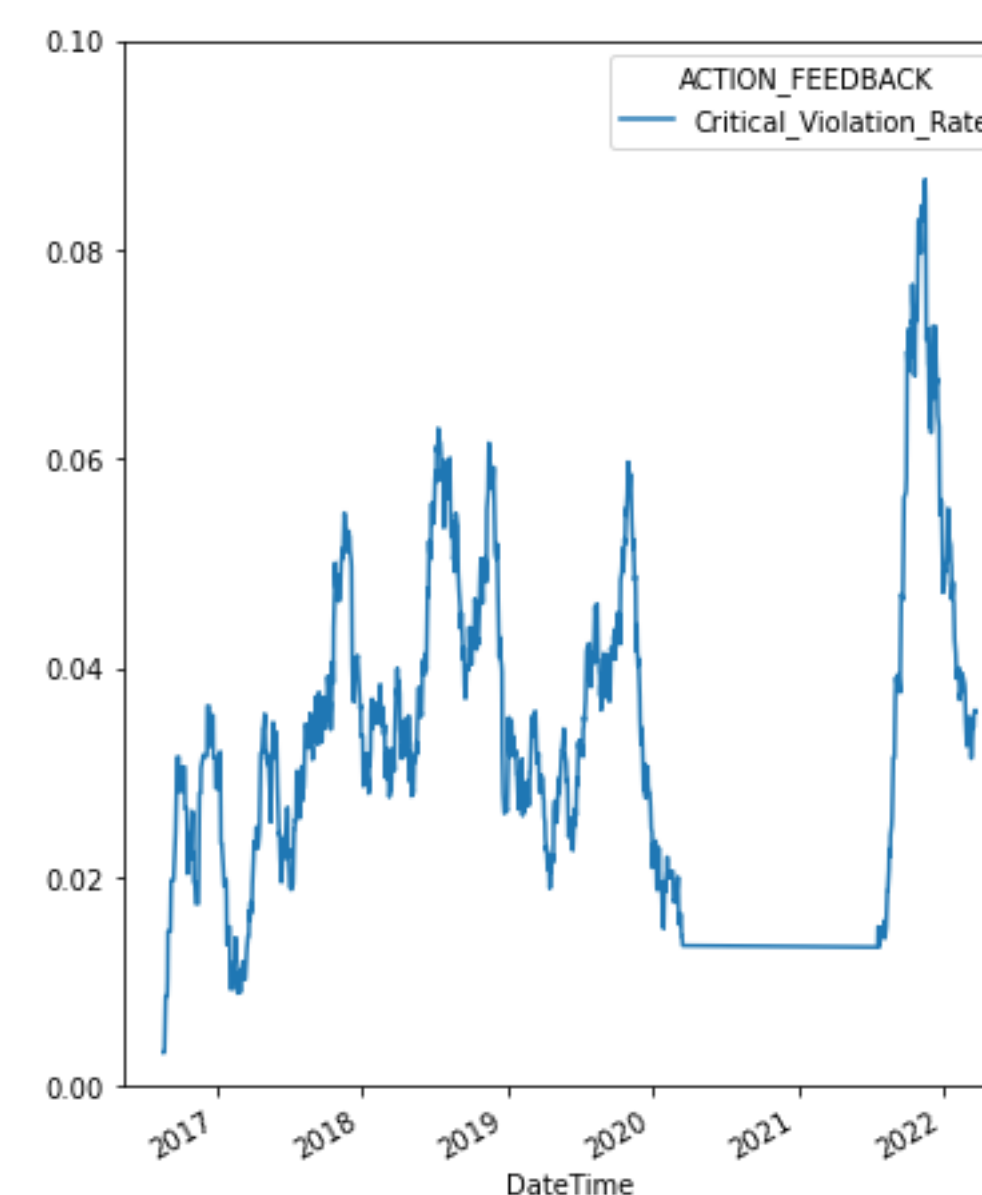
C Grade Word Cloud

We want to know the comments after some restaurants received a C grade, so we extract words from their violation descriptions and then made this word cloud. We can see that there are words relate to annoying creatures and many negative words. It means vermin and flies can be a factor for restaurants to receive a C grade. It seems like avoiding flies and vermin are the keys to getting good grade after the inspections.



Critical Violation Rate Time Series

A small portion of restaurants would receive critical violation warning after inspections. To gain a deeper insight into this issue, we created a time-series plot showing the critical violation rate of the restaurants. After observing the plot, we found that the critical violation rate is not stationary, and the spikes normally happen during the summer and fall seasons. Also, during early-2020 to mid-2021, there were almost no inspection records due to the Covid-19 pandemic.



Interestingly, after the New York City restaurant inspection team started to inspect restaurants again, the critical violation rate reached highest during late 2021. The good news is that the critical violation rate went back to normal in recent few month.

Suggestions

Restaurants should be more careful during summer and fall. These seasons tend to grow vermin and flies, so they should clean their kitchen more often and careful during these times.

Top Six Most Common Restaurant Location Distribution

Here is the six most common restaurant location distribution map. We want to know if some specific types of restaurant would appear in one area. From this map, one thing we can clearly tell is that Japanese restaurant is heavily populated in Manhattan area.

