1. Introduction

Background

In Japan, there are 2 major cities with many different kinds of restaurant. We would like to investigate what is the different of two cities. Sapporo is located in North, Kyoto is located in the middle of Japan.

Based on this study, restaurant investor or commercial ads company can know more about the area and benefit to their business by leverage those insight from the study. For example, if we can know some restaurant type is popular in both cities, then it may be a great opportunity to open chain-restaurant in Japan.

Problem

1. What is the most popular type of restaurant in the area (Kyoto and Sapporo)

2. How strong the competition for the same kind of restaurant by clustering

3. What is the similarities and dissimilarities between Kyoto and Sapporo

1. Data Description

Restaurant Name, Type and Location Coordinates will be expected in JSON. We will work on Kyoto first then go for Sapporo.

Those data are structure data and can be retrieved by using foursquare API.

Kyoto Coordinates: 34.9858,135.7588

Sapporo Coordinates: 43.0678,141.3504



Raw Data from API will be transformed and filtered into data frame. Then, add data attributes for summary. Multiple data frame can be merged to form a summarized information. From the figure above, we cannot directly identify the restaurant belongs to which region (Sapporo or Kyoto).

After that, we can use that information for comparison such as Local Top 10 Restaurant Type, Total Number of Restaurant by Restaurant Type, etc

1. Methodology

Firstly, we will use foursquare API to retrieve data. Entering the parameter including location coordinates, data type (‘Food’) and credentials the it will return the data in JSON format.

Next, we can normalize the json format before load into data frame. Two data frames store the restaurant information by each city. Restaurant information have meta-data and we only need part of the data. New data frame will be used to store organized, aggregated information, it should be enough clear for data visualization.

After that, we can use Folium to visualize the data. Loop through the new data frame and pin point all restaurant into the map by each city.

We can also group those restaurants by different cluster using K-means algorithm. In this case, we try to cluster those restaurants into 5 cluster by city.

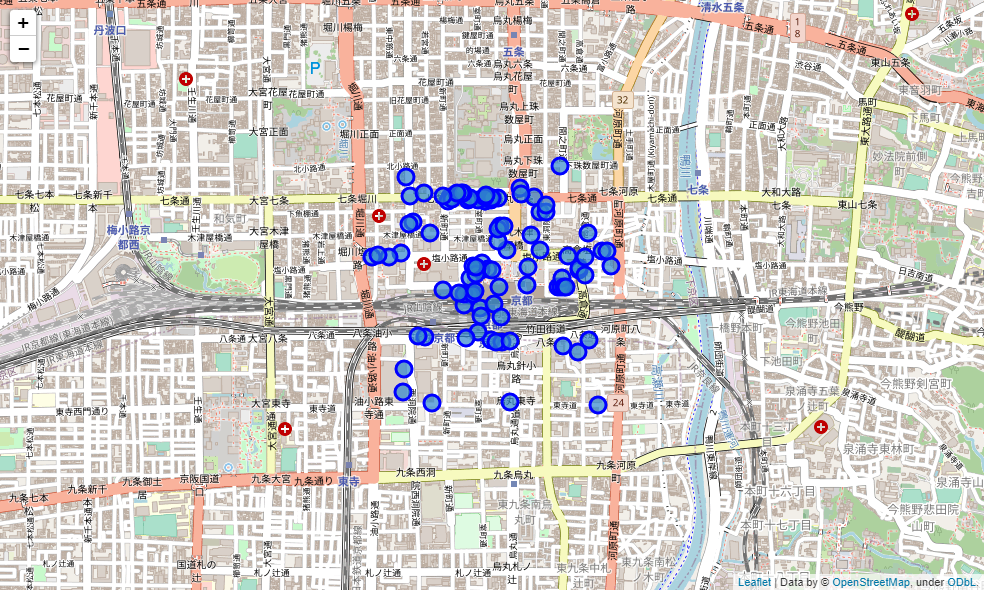
Finally, we have enough information for discussion and answer the problem.

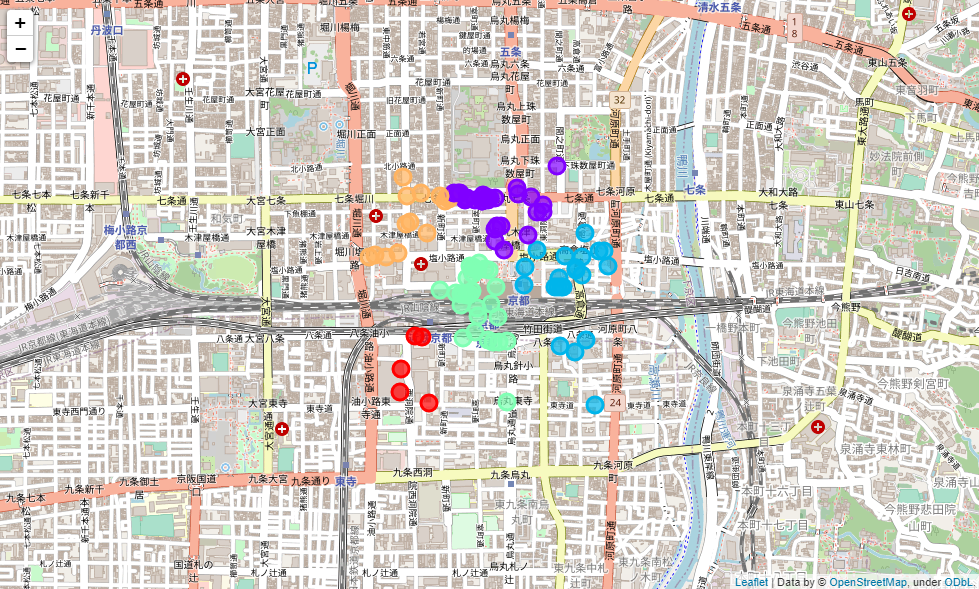
1. Result

Kyoto Restaurant Information (First 5 Rows)



Overview – Restaurant in Kyoto





Cluster 1

Cluster 2

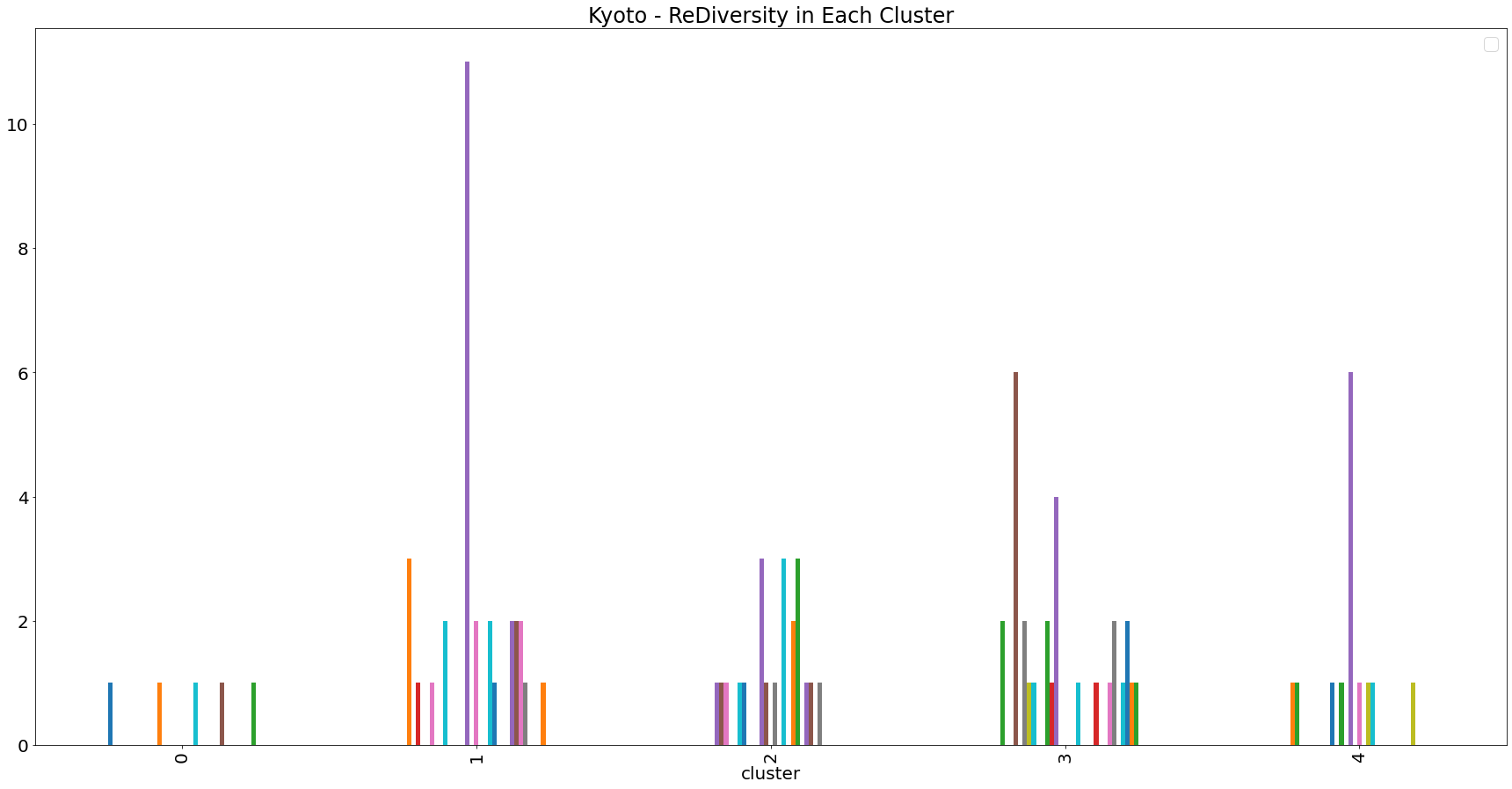
Cluster 3

Cluster 4

Cluster 0

We can see the most of restaurant are located near the train station. Some of them are located near the main street.

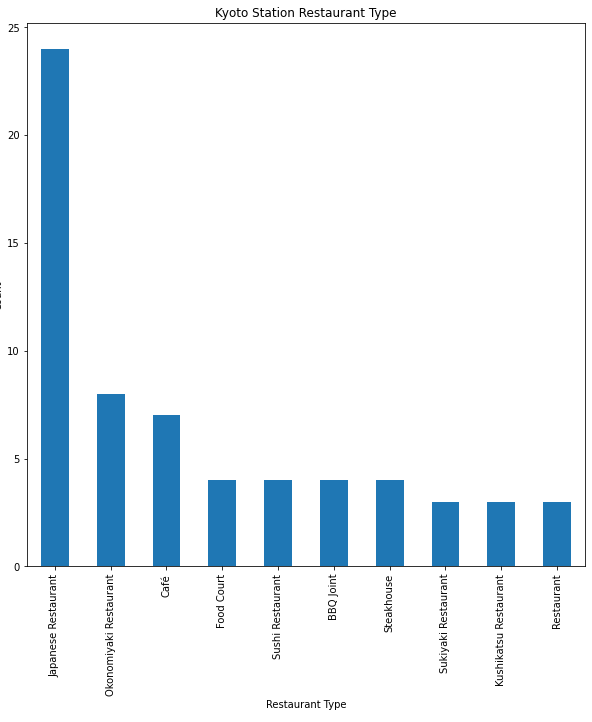
After k-means clustering, we can see the restaurant is clustered into upper-left, north, right, bottom-left and center of Kyoto train station.



|  |  |
| --- | --- |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\305F7157.tmp | C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\D7DE7C7D.tmp |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\ED0E8093.tmp | C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\A091B19.tmp |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\1251458F.tmp |  |

Cluster 0 has less diversity of restaurant type and not much restaurant located in this cluster while Japanese Restaurant is the most popular restaurant type in Cluster 1 and 4.

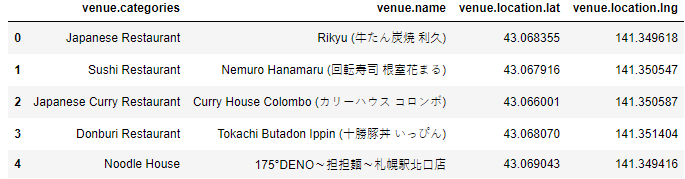
Most of the restaurant in Cluster 3 is café.



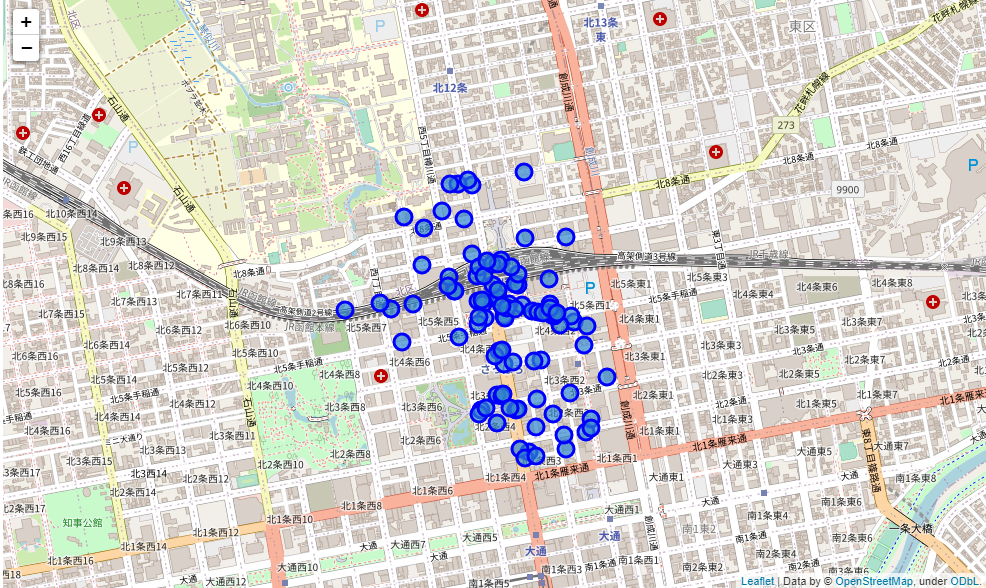
For Kyoto, Japanese Restaurant is the most popular restaurant type, Okonomiyaki Restaurant and Café are 2nd and 3rd most popular.

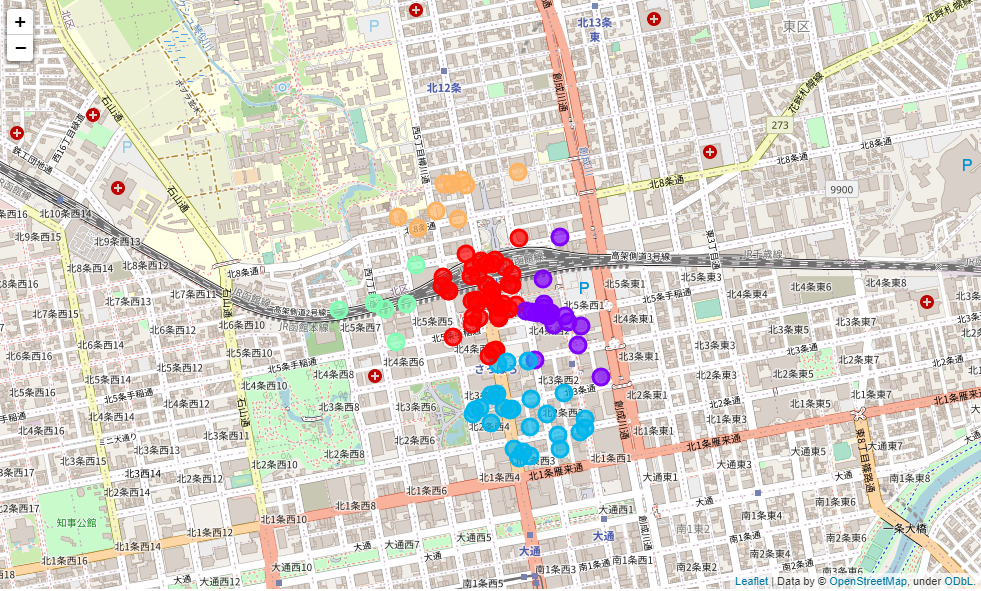
There are more than 20 Japanese restaurant in Kyoto, nearly occupied one-firth of the total.

Sapporo Restaurant Information (First 5 Rows)



Overview – Restaurant in Sapporo





Cluster 2

Cluster 4

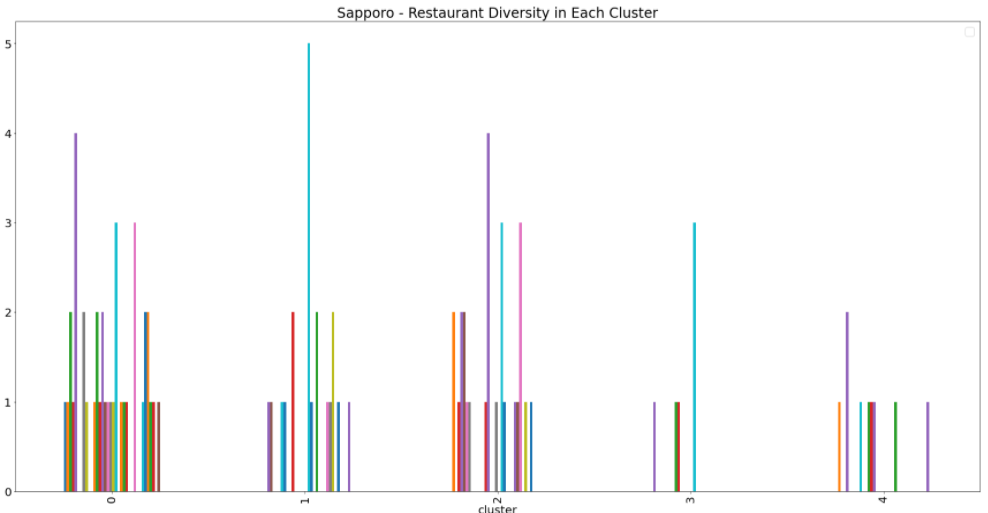
Cluster 3

Cluster 0

Cluster 1

We can see the most of restaurant are located near south side of the train station.

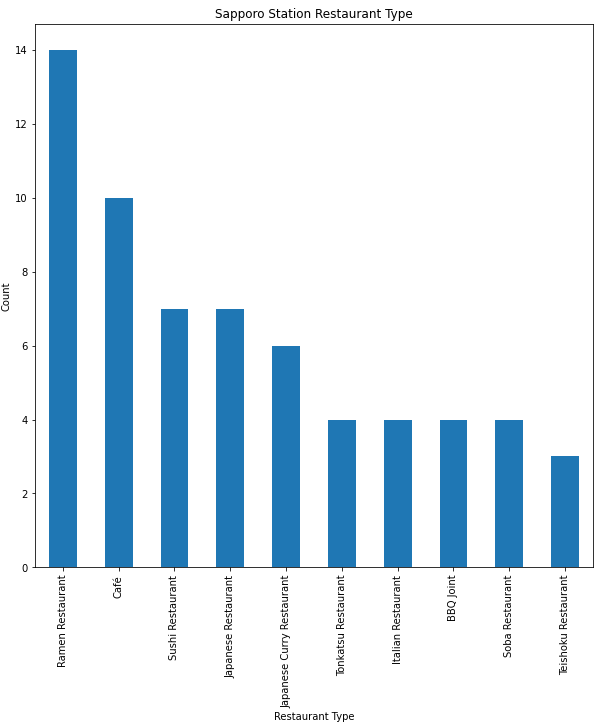
After k-means clustering, we can see the restaurant is clustered into left, north, right, south and center of Sapporo train station.



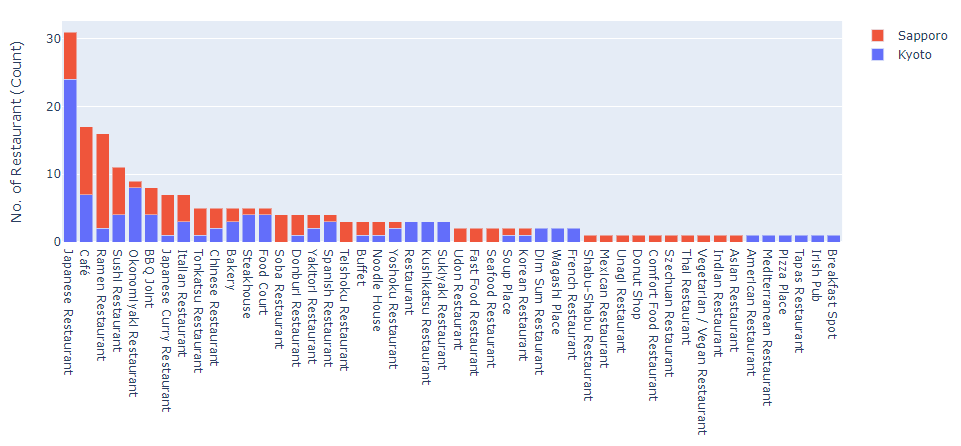
|  |  |
| --- | --- |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\5D07B575.tmp | C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\63FE3C4B.tmp |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\4E022791.tmp | C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\F475A0C7.tmp |
| C:\Users\T07195GESD\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\B4100D6D.tmp |  |

Cluster 0, 1 and 2 are the majority as most of the restaurant located in those cluster. Only some restaurants located in cluster 3 and 4.

Cluster 0 have different type of restaurant and most popular type of restaurant is Café.



Ramen Restaurant is the most popular type in Sapporo. Café and Sushi Restaurant/ Japanese Restaurant are 2nd and 3rd most popular.



Combining two cities restaurant information, we can see the most popular restaurant type is Japanese Restaurant, more than 30 restaurants.

The 2nd most popular restaurant type is Café. More than 15 restaurant in total.

There are several restaurant types that cannot be found in both cities. For example, Dim Sum Restaurant, Irish Pub, Sukiyaki restaurant. Szechuan Restaurant can be found in Sapporo only while Tapas Restaurant only in Kyoto.

1. Discussion

Look like there are too many Japanese Restaurant in both cities. Other restaurant type instead of Japanese Restaurant could be an opportunity to develop in both cities.

I would recommend try on the common restaurant type that can be found in both cities such as Sushi Restaurant, Italian Restaurant, Yakitori Restaurant, BBQ Joint, etc. Those kinds of restaurant have almost 50%-50% of the ratio in both cities. This would be great opportunity for a chain-restaurant investor.

The most competitive area is the center of the train station as the traffic is hot in those are. So, I recommend cluster 0 in Kyoto and cluster 4 in Sapporo for new start-up restaurant. These cluster can have less competition but have less traffic. The land value should be lower than other clusters.

1. Conclusion

Though this study is rough and not very deep-dive but we can still simply know

how the situation by using simple restaurant information.

There are possible future works to improve this study, for example, evaluation of the modelling by measure the accuracy.

It could be developed into national by doing more research in other city as well, then we can have the summary of all restaurant in japan.