# Where to open a new Italian restaurant in Central Yokohama, Naka-ward?

Kaz Watanabe February 6, 2021

#### Introduction

Restaurant business is easy to start and can be profitable when it gets many customers. I have a client who spent long time in Venice, Italy, to run small Italian cafe/restaurant as a chef who is thinking to open a new Italian restaurant in Yokohama because the client thinks the city has enough customers for good Italian foods but misses good Italian restaurants.

### **Business Problem**

I, as a business consultant, agree that Yokohama has lots of potential for Italian restaurant business, but the city is large, and the location is very important to the restaurant business. Within Yokohama, Central Yokohama, especially Naka-ward, is one of the most commercially developed wards in the city, thus it is considered attractive area to open a new Italian restaurant. Yet, Naka-ward has 21 km2 (= square kilometer) with more than 150 thousand people living in the area and day time population becomes 240 thousands people with tourists and office workers visiting, a large and dense area. Within Naka-ward, there are several commercial districts, and it is important to understand their characteristics, and even which Town in district is a better location for Italian restaurant.

### **Data**

- 1. Japan Postal Office Postal Code database: Postal Code is an effective way to segment District, which equates a group of Towns in this case, of the Nakaward. There are 106 Postal codes or Towns in Naka-ward. Postal Code is provided by the Japan Postal office in CSV format.
- Tree-maps database Tree-maps database is a web base geo-coding service used here to associate geographical coordinate with Postal code, so that each Postal Code or Town can have geographical coordinate that then work well with geographical mapping service, Folium, and area characteristic data base, Foursquare.
- Foursuare API: Foursquare is a location-based social networking website, software for mobile devices, and game. Users "check-in" at venues using text messaging or a device specific application. We use its API to obtain Venues and their Category to characterize the area.
- 4. Folium: Folium builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the leaflet.js library. Manipulate the data in Python, then visualize it in on a Leaflet map via folium. This visualization data and function are used to understand density and orientation of Towns in Cluster and Venues on Naka-ward map.

## Methodology

Identify best District within Naka-ward

- 1. Segment Naka-ward by Town: using Japan Post Office postal code to segment the ward into multiple Towns
- 2. Attach Coordinates to Town with Tree-maps database function for Folium map visualization and Foursquare venue data download
- 3. Map Towns over Naka-ward geography with Folium
- 4. Cluster Towns rolling up to District: Use Foursquare venue data to cluster Towns to District by k-mean method
- 5. Review Districts to understand their characteristics and name them for easier reference
- 6. Identify current Italian Restaurants in Naka-ward using Foursquare

#### 1. Segment Naka-ward by Town: using Japan Post Office postal code

Download the Town postal code list of Naka-ward from Post Office web site to local CSV file (= Town.csv)

https://api.nipponsoft.co.jp/zipcode/%E7%A5%9E%E5%A5%88%E5%B7%9D%E7%9C%8C%E6%A8%AA%E6%B5%9C%E5%B8%82%E4%B8%AD%E5%8C%BA

Use Geo-coding web service by uploading the CSV file to get latitude and longitude coordinates by postal code

https://www.tree-maps.com/zip-code-to-coordinate/

Copy the coordinates and paste onto local CSV file (= Coordinate.csv)

Here is the first 5 rows of the table downloaded from Japan Postal Office web site

	Postal_Code	Pref	Ward	Town	Address
0	231-0001	神奈川県	横浜市中区	新港	神奈川県横浜市中区新港
1	231-0002	神奈川県	横浜市中区	海岸通	神奈川県横浜市中区海岸通
2	231-0003	神奈川県	横浜市中区	北仲通	神奈川県横浜市中区北仲通
3	231-0004	神奈川県	横浜市中区	元浜町	神奈川県横浜市中区元浜町
4	231-0005	神奈川県	横浜市中区	本町	神奈川県横浜市中区本町

## 2. Attach Coordinates to Town with Tree-map database function for Folium map visualization and Foursquare venue data download¶

Go to the tree-maps site and download the coordinates data for the postal codes we have, then import local CSV file, Coordinate.csv, to pandas dataframe on the notebook.

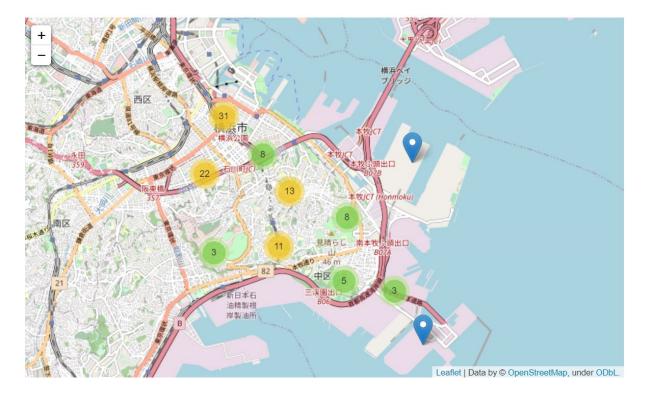
After cleansing the data and join the coordinates data to Postal Codes in Pandas data frame, the table now looks like below.

	Postal_Code	Pref	Ward	Town	Address	Latitude	Longitude
0	231-0001	神奈川県	横浜市中区	新港	神奈川県横浜市中区新港	35.454370	139.641190
1	231-0002	神奈川県	横浜市中区	海岸通	神奈川県横浜市中区海岸通	35.450641	139.642730
2	231-0003	神奈川県	横浜市中区	北仲通	神奈川県横浜市中区北仲通	35.449926	139.637674
3	231-0004	神奈川県	横浜市中区	元浜町	神奈川県横浜市中区元浜町	35.449272	139.640069
4	231-0005	神奈川県	横浜市中区	本町	神奈川県横浜市中区本町	35.449386	139.637551
101	231-0864	神奈川県	横浜市中区	千代崎町	神奈川県横浜市中区千代崎町	35.434236	139.655064
102	231-0865	神奈川県	横浜市中区	北方町	神奈川県横浜市中区北方町	35.433637	139.658545
103	231-0866	神奈川県	横浜市中区	柏葉	神奈川県横浜市中区柏葉	35.432117	139.643034
104	231-0867	神奈川県	横浜市中区	打越	神奈川県横浜市中区打越	35.434222	139.637952
105	231-0868	神奈川県	横浜市中区	石川町	神奈川県横浜市中区石川町	35.436849	139.641109

106 rows × 7 columns

### 3. Map Towns over Naka-ward geography with Folium¶

Using Folium, we can visualize where those Towns are across Naka-ward. Total 106 towns are spread out as below on Naka-ward map. It seems more towns are concentrated on the upper left side of Naka-ward.



## 4. Cluster Towns rolling up to District: Use Foursquare venue data to cluster Towns to District by k-mean method

Since there are too many towns like 106 towns, it would be much easier to cluster those towns in more manageable numbers and identify which cluster or District is more suitable for Italian restaurant after understanding of those Districts.

First step is to access Foursquare API to get venue data of each Town. Using Client ID and Client Secret and assigning the latest version, we get the venue data of 500m radius from each Town center coordinate. The number of venues from a coordinate is limited to 500 each. Total 5,079 rows are retrieved from Foursquare like below.

	Town	Town_Latitude	Town_Longitude	Venue	Venue_Lat	Venue_Long	Venue_Category
0	新港	35.454370	139.641190	Yokohama Hammerhead (横浜ハンマーヘッド)	35.455791	139.641823	Shopping Mall
1	新港	35.454370	139.641190	Port Terrace Cafe	35.454525	139.640704	Café
2	新港	35.454370	139.641190	Yokohama Red Brick Warehouse (横浜赤レンガ 倉庫)	35.452296	139.642874	Historic Site
3	新港	35.454370	139.641190	Akarenga Park (赤レンガパ ーク)	35.454207	139.643201	Park
4	新港	35.454370	139.641190	Granny Smith Apple Pie & Coffee	35.452569	139.642733	Pie Shop
5074	石川 町	35.436849	139.641109	プリンセスガーデン ヨコハ マ	35.440429	139.639518	Bridal Shop
5075	石川 町	35.436849	139.641109	QUO VADIS	35.437902	139.645795	Italian Restaurant
5076	石川 町	35.436849	139.641109	柏葉バス停	35.432482	139.641809	Bus Stop
5077	石川 町	35.436849	139.641109	forgame	35.432570	139.639677	Shoe Store
5078	石川 町	35.436849	139.641109	山元町交差点	35.432632	139.639411	Intersection

5079 rows × 7 columns

This venue data is re-organized to count by Town in below table. This is to check if enough venue data was retrieved by Town. Although there are some towns with small number of venues, we found enough venues in populous towns as expected. The table shows there are 211 unique categories in Naka-ward venues.

	Town_Latitude	Town_Longitude	Venue	Venue_Lat	Venue_Long	Venue_Category
Town						
かもめ町	3	3	3	3	3	3
万代町	60	60	60	60	60	60
三吉町	32	32	32	32	32	32
上野町	22	22	22	22	22	22
不老町	52	52	52	52	52	52
野毛町	100	100	100	100	100	100
長者町	72	72	72	72	72	72
鷺山	14	14	14	14	14	14
麦田町	14	14	14	14	14	14
黄金町	37	37	37	37	37	37

<sup>105</sup> rows × 6 columns

## 4. Cluster Towns rolling up to District (continued) with k-mean method

In [22]:

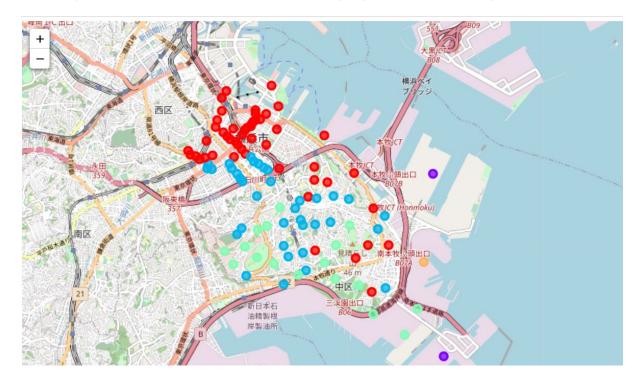
Using Venue data of each Town, we will run Cluster analysis with k-mean method. First thing to do is to assess the mean of occurrence of each category by Town and sort the categories of venue by popularity (= frequency of occurrence) in each Town.

```
----本町----
               venue freq
0
               Café 0.06
        Coffee Shop 0.05
  Italian Restaurant 0.05
     Soba Restaurant 0.04
4 Japanese Restaurant 0.04
----本郷町----
          venue freq
   Convenience Store 0.24
    Bus Stop 0.10
2 Chinese Restaurant 0.10
           Bakery 0.10
      Grocery Store 0.05
----松影町----
              venue freq
O Convenience Store 0.18
    Ramen Restaurant 0.08
    Baseball Stadium 0.08
    Soba Restaurant 0.04
4 Italian Restaurant 0.04
----柏葉----
             venue freq
0 Bus Stop 0.2.
1 Historic Site 0.09
```

After rearranging the data frame for k-mean method, we ran k-mean method with different number of clusters. We landed 5 is the number of clusters appropriate in this analysis, and below table is the excerpt of the result table. The column "Cluster Labels" now has values from 0.0 to 4.0 applied to each Town row.

	Postal_Code	Town	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5t Cc
0	231-0001	新港	35.454370	139.641190	0.0	Café	Shopping Mall	Seafood Restaurant	Convenience Store	Italian Resta
1	231-0002	海岸通	35.450641	139.642730	0.0	Café	History Museum	Historic Site	Clothing Store	Conve Store
2	231-0003	北仲 通	35.449926	139.637674	0.0	Café	Italian Restaurant	Hotel	Bed & Breakfast	Japan Resta
3	231-0004	元浜 町	35.449272	139.640069	0.0	Café	Hotel	History Museum	Italian Restaurant	Jazz (
4	231-0005	本町	35.449386	139.637551	0.0	Café	Italian Restaurant	Coffee Shop	Soba Restaurant	Hotel
5	231-0006	南仲 通	35.448198	139.638584	0.0	Café	Coffee Shop	Bed & Breakfast	Convenience Store	Tonka Resta
6	231-0007	弁天 通	35.448037	139.637668	0.0	Café	Coffee Shop	Convenience Store	Bed & Breakfast	Japan Resta

Since the number of rows is still 106 and the table does not give us a good glimpse of clustering result, we should map the towns by cluster over Naka-ward. Below is the result of Town clustering, where Red is 0, Purple is 1, Blue is 2, Light green is 3, and Orange is 4.



## 5. Review Districts to understand their characteristics and name them for easier reference

In order to understand the characteristics of each cluster, we need to see Most Common Venues in each cluster. Assessing through them per cluster, we named each cluster as District according to the natures of Most Common Venues.

Cluster 0: Business & Commercial District

	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	新港	Café	Shopping Mall	Seafood Restaurant	Convenience Store	Italian Restaurant	Hotel	Coffee Shop	Park	Japanese Restaurant	Mexican Restaurant
1	海岸 通	Café	History Museum	Historic Site	Clothing Store	Convenience Store	Shopping Mall	Italian Restaurant	Pie Shop	Hawaiian Restaurant	American Restaurant
2	北仲 通	Café	Italian Restaurant	Hotel	Bed & Breakfast	Japanese Restaurant	Tonkatsu Restaurant	History Museum	Sake Bar	Coffee Shop	Soba Restaurant
3	元浜 町	Café	Hotel	History Museum	Italian Restaurant	Jazz Club	BBQ Joint	Sake Bar	Coffee Shop	Convenience Store	Bed & Breakfast
4	本町	Café	Italian Restaurant	Coffee Shop	Soba Restaurant	Hotel	Bed & Breakfast	Japanese Restaurant	Sake Bar	History Museum	Tonkatsu Restaurant
5	南仲 通	Café	Coffee Shop	Bed & Breakfast	Convenience Store	Tonkatsu Restaurant	History Museum	Japanese Restaurant	Museum	Park	Soba Restaurant

Cluster 1	: Shop &	Entertai	inment Dis	strict
		2nd Most	3rd Most	4th Mo

	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
18	松影町	Convenience Store	Ramen Restaurant	Baseball Stadium	Historic Site	Soba Restaurant	Bed & Breakfast	Intersection	Italian Restaurant	Japanese Restaurant	Beer Garden
19	寿町	Convenience Store	Ramen Restaurant	Baseball Stadium	Chinese Restaurant	Coffee Shop	Sporting Goods Shop	Café	Soba Restaurant	Bed & Breakfast	Bar
20	扇町	Convenience Store	Baseball Stadium	Ramen Restaurant	Grocery Store	Soba Restaurant	Bed & Breakfast	Intersection	Sporting Goods Shop	Donburi Restaurant	Café
21	翁町	Convenience Store	Baseball Stadium	Soba Restaurant	Ramen Restaurant	Grocery Store	Rock Club	Intersection	Japanese Restaurant	Sporting Goods Shop	Donburi Restaurant
23	不老町	Convenience Store	Ramen Restaurant	Baseball Stadium	Soba Restaurant	Café	Sake Bar	Rock Club	Drugstore	Sporting Goods Shop	Soup Place
25	三吉町	Convenience Store	Ramen Restaurant	Japanese Family Restaurant	Grocery Store	Donburi Restaurant	Bar	Park	Wine Bar	Comedy Club	Bus Station

#### Cluster 2: Port District

То	own	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
	本牧 禄ケ 丘	Sports Club	Park	Playground	Zoo Exhibit	Fried Chicken Joint	Fountain	Food Court	Food & Drink Shop	Food	Flower Shop
74 池	也袋	Park	Sports Club	Convenience Store	Ramen Restaurant	Electronics Store	French Restaurant	Fountain	Food Court	Food & Drink Shop	Food
	限岸 加曽 台	Sports Club	Park	Bus Stop	Convenience Store	Zoo Exhibit	Event Space	French Restaurant	Fountain	Food Court	Food & Drink Shop

Cluster 3: Park and Residential District

	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
57	本牧町	Bus Stop	Convenience Store	Park	Grocery Store	Szechuan Restaurant	Bakery	Pharmacy	Chinese Restaurant	Snack Place	Bar
60	かも め町	Intersection	Bus Station	Bus Stop	Food & Drink Shop	Zoo Exhibit	Fast Food Restaurant	Garden	Furniture / Home Store	Fried Chicken Joint	French Restaurant
61	豊浦町	Train Station	Intersection	Bus Stop	Event Space	Furniture / Home Store	Fried Chicken Joint	French Restaurant	Fountain	Food Court	Food & Drink Shop
62	千鳥 町	Park	Tennis Court	Bus Stop	Toll Booth	Pool	Historic Site	History Museum	Baseball Field	Bus Station	Intersection
67	本牧 三之 谷	Historic Site	Park	Bus Stop	Convenience Store	Bakery	Garden	Café	Tea Room	Grocery Store	Snack Place

#### Cluster 4: Industrial District

	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
61	豊浦町	Train Station	Intersection	Bus Stop	Zoo Exhibit	French Restaurant	Fountain	Food Court	Food & Drink Shop	Food	Flower Shop
62	千鳥 町	Park	Bus Stop	Tennis Court	Baseball Field	Historic Site	History Museum	Pool	Toll Booth	Bus Station	Intersection
65	本牧 元町	Bus Station	Bus Stop	Convenience Store	Intersection	History Museum	Park	Event Space	French Restaurant	Fountain	Food Court
66	本牧 大里 町	Historic Site	Park	Garden	Intersection	Pool	Snack Place	Bus Stop	Bus Station	Tea Room	Tennis Court
67	本牧 三之 谷	Historic Site	Bus Stop	Park	Convenience Store	Garden	Bakery	Liquor Store	Clothing Store	Tea Room	History Museum

From this assessment, the first two Districts are more suited for new Italian Restaurant, as they have more restaurants in Most Common Venues.

## 6. Identify Italian Restaurants in Cluster Labels 0 or 1 Districts using Foursquare

Since we now know the two best Districts for new Italian restaurant, we can narrow down our analysis into those Districts. First, let's see where and how many Italian Restaurants we already have in the Districts. We can retrieve the Italian restaurant venues from Foursquare and map them by Town and District. Below table shows excerpt of the Italian Restaurant counts by Town.

	Town_Latitude	Town_Longitude	Venue	Venue_Lat	Venue_Long	Venue_Category	CI
Town							
万代町	1	1	1	1	1	1	
不老町	1	1	1	1	1	1	
住吉町	5	5	5	5	5	5	
元浜町	2	2	2	2	2	2	
内田町	1	1	1	1	1	1	
北仲通	3	3	3	3	3	3	
南仲通	3	3	3	3	3	3	
吉田町	5	5	5	5	5	5	
太田町	5	5	5	5	5	5	
宮川町	6	6	6	6	6	6	

Also we should map those Italian Restaurants over Naka-ward to see where they are. Below is the map with Folium.



## **Results**

There is fair concentration of restaurants in a couple of Districts rather than spreading out over entire Naka-ward. Cluster analysis identified 2 out of 5 Districts having many restaurants, but when focusing just on the category "Italian Restaurant", it is only 1 District, Business & Commercial District, where we can see many Italian restaurants exist today

	Postal_Code	Town	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	
0	231-0001	新港	35.454370	139.641190	0.0	Café	Shopping Mall	Seafood Restaurant	Convenience Store	Italian Restaurant	Hotel	Coffee Shop	Park	
1	231-0002	海岸通	35.450641	139.642730	0.0	Café	History Museum	Historic Site	Clothing Store	Convenience Store	Shopping Mall	Italian Restaurant	Pie Shop	
2	231-0003	北仲 通	35.449926	139.637674	0.0	Café	Italian Restaurant	Hotel	Bed & Breakfast	Japanese Restaurant	Tonkatsu Restaurant	History Museum	Sake Bar	(
3	231-0004	元浜 町	35.449272	139.640069	0.0	Café	Hotel	History Museum	Italian Restaurant	Jazz Club	BBQ Joint	Sake Bar	Coffee Shop	C
4	231-0005	本町	35.449386	139.637551	0.0	Café	Italian Restaurant	Coffee Shop	Soba Restaurant	Hotel	Bed & Breakfast	Japanese Restaurant	Sake Bar	

### **Discussion**

Even in Business & Commercial District, there are two Towns listing Italian Restaurant as 2nd Most Common Venue. We should first review such Towns if the area has appropriate atmosphere. Recommend either Postal Code 231-0003 or 0005 , "北中通" or "本町", to first consider the Town to open new Italian Restaurant. Those two towns appear to have many Cafe and Hotel without distractive shops which combined offer more attractive atmosphere for Italian restaurant.

Comparing the two towns with area photos, they are two very different towns. "北中通" or Kitanaka is on the left photo, and it is a modern business district with offices and shops in the buildings. New Italian restaurant will be in one of the high raised buildings and on restaurant floor. On the other hand, "本町" or Honmachi, on the right photo, has open and wide promenade with many tourists all the time. New Italian restaurant will preferably be on the street and appealing to tourist traffic.





## Conclusion

This analysis identified the two most preferable towns for new Italian restaurant, but the towns have two very different environment and dynamics. Next step is to develop restaurant concept(s) which would be much dependent on which town it opens in, as they target two different customers and will be surrounded by two different atmospheres. Lots of details have to be considered for concept development, but at least we have now two very best towns in Central Yokohama, Naka-ward, as recommendation..!

[end of memo]