

The Optimal Location to Open a Chinese Restaurant in Tokyo, Japan

Cherrie He

June 10, 2020

Introduction

Background

To be successful with any restaurant opening, market research is the best thing to start off.

At one time, market researches before running a business were implemented in traditional ways, like visiting local communities or interviewing potential customers. Whereas nowadays, the researches are largely facilitated since the machine learning power were added.

Data analysis and machine learning techniques play very important role in helping business people grab market share and become profitable when opening a restaurant in a smart and scientific way.

Business Problem

A client seeks to establish a Chinese restaurant in Tokyo, Japan.

This capstone report will determine the optimal and most strategic location for running the restaurant by exploring the following questions:

- Is the market saturated?
- Who are the potential customers?
- Who are the local competitors?

Interests

This report will be targeted to stakeholders interested in opening an Chinese Cuisine restaurant in Tokyo, Japan. Various factors such as over-saturation or no demand ,for the type of restaurant that the customer wants to open, effect the success or failure of the restaurant. Hence, customers can bolster their decisions using the descriptive and predictive capabilities of data science.

Data Acquisition and Wrangling

Data Sources

Tokyo Wards Table from Wikipedia

I first make use of Special Wards of Tokyo page from [Wikipedia - Special Wards of Tokyo](#) to scrap the table to create a data-frame. Simply use pandas.read_html and then extract the target table.

Tokyo land price are downloaded from [Tokyo land market value](#)

For the demographical data, which will help analyse the potential customers are downloaded from [TOKYO STATISTICAL YEARBOOK 2018](#)

	Name	Population	Density	Area	Major district
0	Chiyoda	59441	5100	11.66	Nagatacho
1	Chuo	147620	14460	10.21	Nihonbashi
2	Minato	248071	12180	20.37	Odaiba
3	Shinjuku	339211	18620	18.22	Shinjuku
4	Bunkyo	223389	19790	11.29	Hongo

As we can see, there are too many districts major to each ward. Here I keep the first one as the major district of the ward.

	Name	Population	Density	Area	Major district
13	Nakano	332902	21350	15.59	Nakano
14	Suginami	570483	16750	34.06	Koenji
15	Toshima	294673	22650	13.01	Ikebukuro
16	Kita	345063	16740	20.61	Akabane
17	Arakawa	213648	21030	10.16	Arakawa
18	Itabashi	569225	17670	32.22	Itabashi
19	Nerima	726748	15120	48.08	Nerima
20	Adachi	674067	12660	53.25	Ayase
21	Katsushika	447140	12850	34.80	Tateishi
22	Edogawa	685899	13750	49.90	Kasai

Tokyo Land Prices and Demographical Feature Data

	Major District	Average Unit Price(₹/sq.m)	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	India	Myanmar	Thailand	Others	Total
0	Chiyoda	2602846.0	1250.0	455.0	76.0	64.0	24.0	210.0	79.0	16.0	46.0	582.0	2996.0
1	Chuo	2688537.0	3266.0	1401.0	192.0	153.0	95.0	391.0	275.0	38.0	89.0	1366.0	7651.0
2	Minato	2180846.0	3962.0	3461.0	144.0	1027.0	107.0	3257.0	649.0	56.0	191.0	6447.0	20057.0
3	Shinjuku	940756.0	14153.0	10221.0	3484.0	747.0	3517.0	1033.0	246.0	2218.0	735.0	4830.0	43068.0
4	Bunkyo	943847.0	4646.0	1658.0	927.0	223.0	364.0	329.0	107.0	317.0	192.0	1536.0	10808.0

Getting the Longitudes and Latitudes of Major Districts from Geopy Client

Next objective is to get the coordinates of these 23 major districts using geocoder class of Geopy client

	Name	Population	Density	Area	Major district	lat	long
0	Chiyoda	59441	5100	11.66	Nagatacho	35.675618	139.743469
1	Chuo	147620	14460	10.21	Nihonbashi	35.684068	139.774503
2	Minato	248071	12180	20.37	Odaiba	35.619050	139.779364
3	Shinjuku	339211	18620	18.22	Shinjuku	35.693763	139.703632
4	Bunkyo	223389	19790	11.29	Hongo	35.175376	137.013476
5	Taito	200486	19830	10.11	Ueno	35.711759	139.777645
6	Sumida	260358	18910	13.77	Kinshicho	35.696312	139.815043
7	Koto	502579	12510	40.16	Kiba	23.013134	-80.832875
8	Shinagawa	392492	17180	22.84	Shinagawa	35.599252	139.738910
9	Meguro	280283	19110	14.67	Meguro	35.621250	139.688014
10	Ota	722608	11910	60.66	Omori	35.588473	139.727933
11	Setagaya	910868	15690	58.05	Setagaya	35.646096	139.656270
12	Shibuya	227850	15080	15.11	Shibuya	35.664596	139.698711
13	Nakano	332902	21350	15.59	Nakano	35.718123	139.664468
14	Suginami	570483	16750	34.06	Koenji	35.704942	139.649909
15	Toshima	294673	22650	13.01	Ikebukuro	35.730103	139.711884
16	Kita	345063	16740	20.61	Akabane	35.778139	139.720800
17	Arakawa	213648	21030	10.16	Arakawa	35.737529	139.781310
18	Itabashi	569225	17670	32.22	Itabashi	35.774143	139.681209
19	Nerima	726748	15120	48.08	Nerima	35.748360	139.638735
20	Adachi	674067	12660	53.25	Ayase	35.446047	139.430823
21	Katsushika	447140	12850	34.80	Tateishi	33.481791	131.478154
22	Edogawa	685899	13750	49.90	Kasai	-5.349800	21.424098

The Percentage of Chinese Residents in All Districts

Some coordinations are not correct, therefore the data needs to be updated.

	Name	Population	Density	Area	Major district	lat	long	Average Unit Price\\$/t(JPY/sq.m)	China	Korea	Vietnam	Philippines
0	Chiyoda	59441	5100	11.66	Nagatacho	35.675618	139.743469	2602846.0	1250.0	455.0	76.0	64.0
1	Chuo	147620	14460	10.21	Nihonbashi	35.684068	139.774503	2688537.0	3266.0	1401.0	192.0	153.0
2	Minato	248071	12180	20.37	Odaiba	35.619050	139.779364	2180846.0	3962.0	3461.0	144.0	1027.0
3	Shinjuku	339211	18620	18.22	Shinjuku	35.693763	139.703632	940756.0	14153.0	10221.0	3484.0	747.0
4	Bunkyo	223389	19790	11.29	Hongo	35.175376	137.013476	943847.0	4646.0	1658.0	927.0	223.0

Location Data by Foursquare API

For returning the popular spots in the vicinity of each neighbourhood, Foursquare API will be utilized, more specifically, its explore function.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Nagatacho	35.675618	139.743469	Nagatacho Kurosawa (永田町 黒澤)	35.674699	139.741737	Japanese Restaurant
1	Nagatacho	35.675618	139.743469	The Capitol Hotel Tokyu (ザ・キャピトルホテル東急)	35.673927	139.741019	Hotel
2	Nagatacho	35.675618	139.743469	Tully's Coffee	35.674594	139.743007	Coffee Shop
3	Nagatacho	35.675618	139.743469	Shinamen Hashigo (支那麺 はしご)	35.672184	139.741576	Ramen Restaurant
4	Nagatacho	35.675618	139.743469	All Day Dining Origami (オールディーダイニング ORIGAMI)	35.673815	139.741104	Restaurant

Methodology

Exploratory Data Analysis

Distributions of Restaurants in Tokyo

Narrow down the venues to the places serve food

By filtering all restaurants and other places serve foods out of the venues dataframe, extract the information of local restaurant distributions. There are 195 unique categories. I will choose the places that contain "Restaurant" or seemingly serve food.

There are 195 venue categories

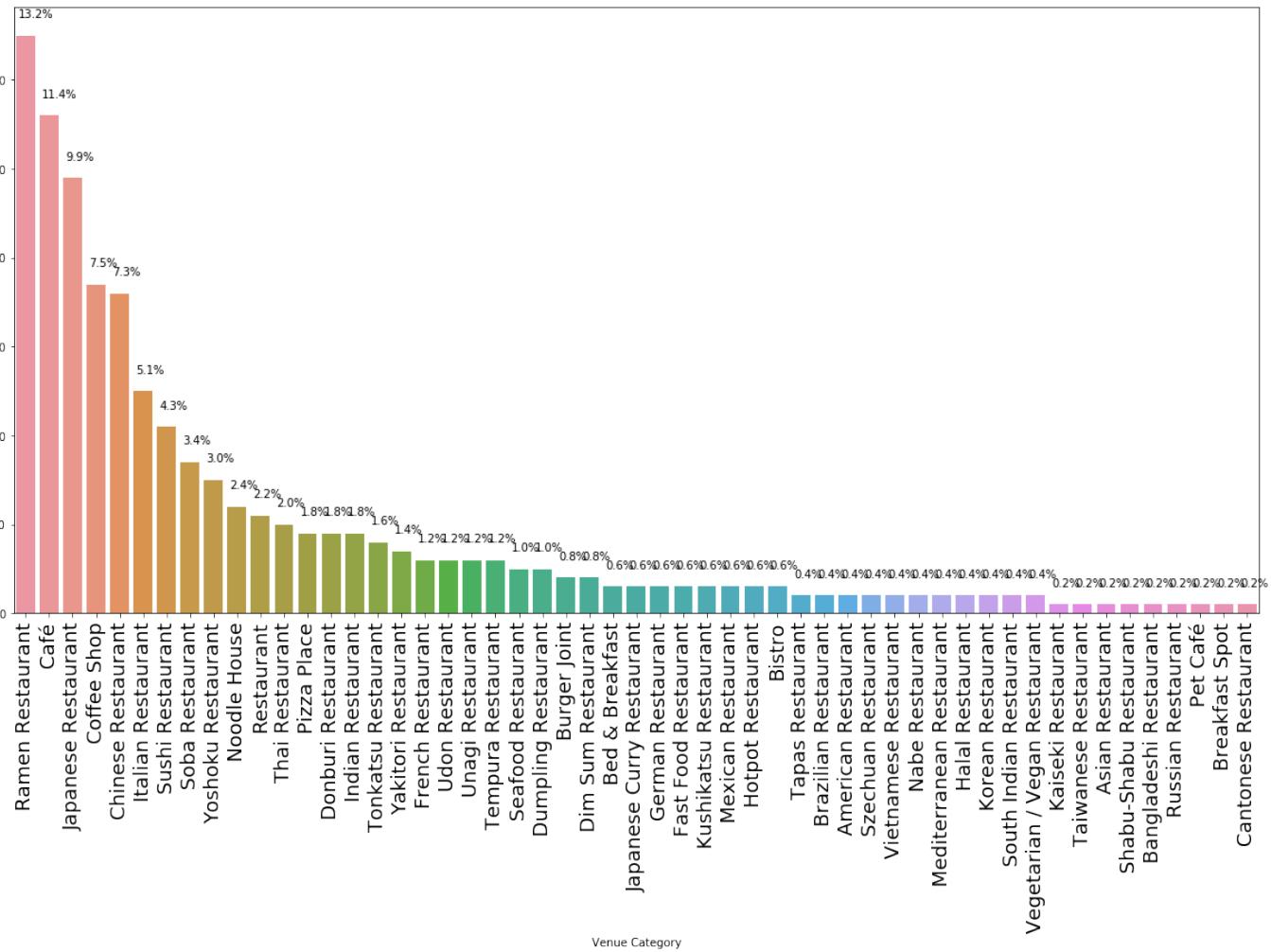
Chinese Restaurant Distributions and Restaurant Ranks

There are 493 out of 1201 nearby venues serve food. By plotting a bar chart and annotating the frequency of their distributions, it is clearly that the most popular restaurant is "Ramen Restaurant" which has the top frequency of distribution at 13.8%. Followed by "Cafe" at 10.8%, "Japanese Restaurant" at 9.1%, "Chinese Restaurant" at 7.5% and "Coffee Shop" at 6.9%.

The number of Chinese restaurants are as half as that of Ramen restaurants, and it is 1% less than Japanese restaurant. Whereas the restaurants ranked below No.5 took at most half of the frequency of Chinese restaurants.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Nagatacho	35.675618	139.743469	Nagatacho Kurosawa (永田町 黒澤)	35.674699	139.741737	Japanese Restaurant
2	Nagatacho	35.675618	139.743469	Tully's Coffee	35.674594	139.743007	Coffee Shop
3	Nagatacho	35.675618	139.743469	Shinamen Hashigo (支那麺 はしご)	35.672184	139.741576	Ramen Restaurant
4	Nagatacho	35.675618	139.743469	All Day Dining Origami (オールディーダイニング ORIGAMI)	35.673815	139.741104	Restaurant
5	Nagatacho	35.675618	139.743469	Sushi Isshin (鮓一新)	35.672589	139.739399	Sushi Restaurant

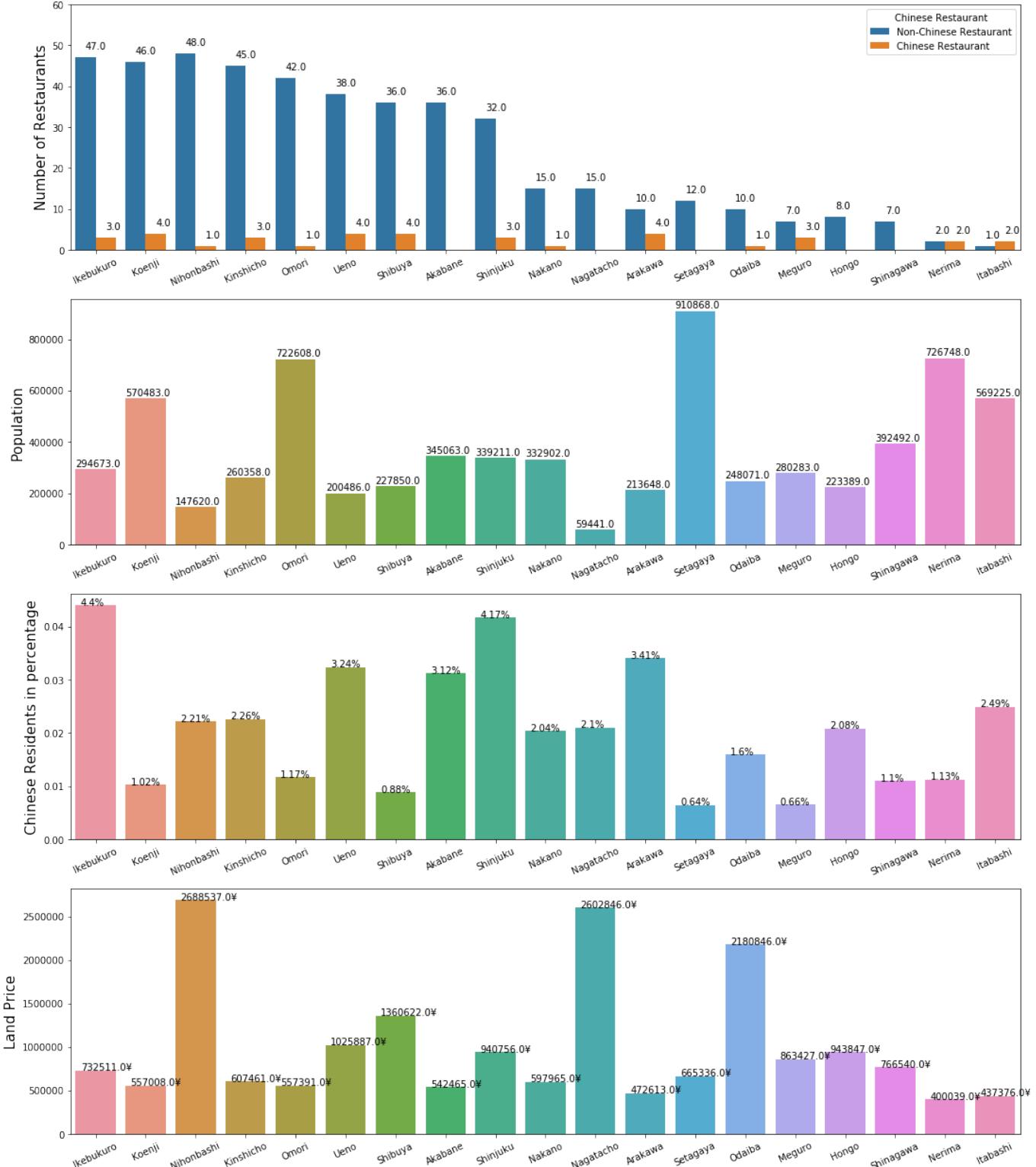
(493, 7)



What about the restaurant distribution of each ward?

The number of restaurants and Tokyo population are presented as follows. This will help analyse the relationship between market requirements and market size.

Number of Restaurants and Population in Each Neighborhood



Neighborhood	Venue	
	Venue Category	
Akabane	Ramen Restaurant	6
	Soba Restaurant	4

	Coffee Shop	3
	Café	2
	Dumpling Restaurant	2
	Indian Restaurant	2
	Italian Restaurant	2
	Japanese Curry Restaurant	2
	Japanese Restaurant	2
	Kushikatsu Restaurant	2
	Yakitori Restaurant	2
	Bistro	1
	Donburi Restaurant	1
	Fast Food Restaurant	1
	Noodle House	1
	Restaurant	1
	Sushi Restaurant	1
	Unagi Restaurant	1
Arakawa	Chinese Restaurant	4
	Ramen Restaurant	3
	Burger Joint	1
	Café	1
	Coffee Shop	1
	Donburi Restaurant	1
	Italian Restaurant	1
	Japanese Restaurant	1
	Noodle House	1
Hongo	Café	2
	Japanese Restaurant	1
	Mediterranean Restaurant	1
	Pizza Place	1
	Ramen Restaurant	1
	Restaurant	1
	Thai Restaurant	1
Ikebukuro	Café	8
	Japanese Restaurant	6
	Sushi Restaurant	6
	Yoshoku Restaurant	5
	Coffee Shop	4
	Ramen Restaurant	4
	Chinese Restaurant	3
	Dim Sum Restaurant	2
	Asian Restaurant	1
	Donburi Restaurant	1

	French Restaurant	1
	German Restaurant	1
	Hotpot Restaurant	1
	Italian Restaurant	1
	Noodle House	1
	Pet Café	1
	Pizza Place	1
	Seafood Restaurant	1
	Szechuan Restaurant	1
	Udon Restaurant	1
Itabashi	Chinese Restaurant	2
	Restaurant	1
Kinshicho	Ramen Restaurant	9
	Coffee Shop	6
	Chinese Restaurant	3
	Thai Restaurant	3
	Japanese Restaurant	2
	Sushi Restaurant	2
	Bangladeshi Restaurant	1
	Café	1
	Cantonese Restaurant	1
	Dumpling Restaurant	1
	Fast Food Restaurant	1
	German Restaurant	1
	Halal Restaurant	1
	Indian Restaurant	1
	Italian Restaurant	1
	Korean Restaurant	1
	Nabe Restaurant	1
	Pizza Place	1
	Restaurant	1
	Seafood Restaurant	1
	Soba Restaurant	1
	South Indian Restaurant	1
	Taiwanese Restaurant	1
	Tonkatsu Restaurant	1
	Udon Restaurant	1
	Unagi Restaurant	1
	Vietnamese Restaurant	1
	Yakitori Restaurant	1
	Yoshoku Restaurant	1
Koenji	Ramen Restaurant	9

	Café	8
	Italian Restaurant	6
	Chinese Restaurant	4
	Dumpling Restaurant	2
	Indian Restaurant	2
	Japanese Restaurant	2
	Mexican Restaurant	2
	Noodle House	2
	Yakitori Restaurant	2
	American Restaurant	1
	Pizza Place	1
	Seafood Restaurant	1
	Soba Restaurant	1
	Sushi Restaurant	1
	Tempura Restaurant	1
	Thai Restaurant	1
	Udon Restaurant	1
	Vegetarian / Vegan Restaurant	1
	Vietnamese Restaurant	1
	Yoshoku Restaurant	1
Meguro	Chinese Restaurant	3
	Coffee Shop	2
	Japanese Restaurant	2
	Italian Restaurant	1
	Noodle House	1
	Sushi Restaurant	1
Nagatacho	Coffee Shop	4
	Japanese Restaurant	3
	Café	1
	Indian Restaurant	1
	Kaiseki Restaurant	1
	Pizza Place	1
	Ramen Restaurant	1
	Restaurant	1
	Soba Restaurant	1
	Sushi Restaurant	1
Nakano	Ramen Restaurant	4
	Japanese Restaurant	3
	Café	2
	Italian Restaurant	2
	Noodle House	2
	Chinese Restaurant	1

	Coffee Shop	1
	Donburi Restaurant	1
Nerima	Café	2
	Chinese Restaurant	2
Nihonbashi	Coffee Shop	6
	Japanese Restaurant	6
	Ramen Restaurant	4
	Yoshoku Restaurant	4
	Café	3
	Italian Restaurant	3
	Dim Sum Restaurant	2
	French Restaurant	2
	Soba Restaurant	2
	Sushi Restaurant	2
	Tempura Restaurant	2
	Bistro	1
	Chinese Restaurant	1
	Donburi Restaurant	1
	German Restaurant	1
	Hotpot Restaurant	1
	Japanese Curry Restaurant	1
	Mediterranean Restaurant	1
	Seafood Restaurant	1
	Tapas Restaurant	1
	Thai Restaurant	1
	Tonkatsu Restaurant	1
	Udon Restaurant	1
	Unagi Restaurant	1
Odaiba	Japanese Restaurant	3
	Coffee Shop	2
	Chinese Restaurant	1
	Italian Restaurant	1
	Ramen Restaurant	1
	Restaurant	1
	Sushi Restaurant	1
	Yakitori Restaurant	1
Omori	Coffee Shop	6
	Café	4
	Ramen Restaurant	4
	Indian Restaurant	3
	Japanese Restaurant	3
	Soba Restaurant	3

	Italian Restaurant	2
	Sushi Restaurant	2
	Tempura Restaurant	2
	Tonkatsu Restaurant	2
	Udon Restaurant	2
	Breakfast Spot	1
	Burger Joint	1
	Chinese Restaurant	1
	French Restaurant	1
	Korean Restaurant	1
	Noodle House	1
	Restaurant	1
	Tapas Restaurant	1
	Thai Restaurant	1
	Yoshoku Restaurant	1
Setagaya	Café	3
	Ramen Restaurant	2
	Bistro	1
	Fast Food Restaurant	1
	Japanese Restaurant	1
	Pizza Place	1
	Restaurant	1
	Szechuan Restaurant	1
	Unagi Restaurant	1
Shibuya	Café	15
	Chinese Restaurant	4
	Japanese Restaurant	4
	Ramen Restaurant	4
	French Restaurant	2
	Brazilian Restaurant	1
	Burger Joint	1
	Coffee Shop	1
	Donburi Restaurant	1
	Italian Restaurant	1
	Mexican Restaurant	1
	Nabe Restaurant	1
	Noodle House	1
	Restaurant	1
	Shabu-Shabu Restaurant	1
	South Indian Restaurant	1
Shinagawa	Donburi Restaurant	2
	Japanese Restaurant	1

	Pizza Place	1
	Restaurant	1
	Soba Restaurant	1
	Sushi Restaurant	1
Shinjuku	Ramen Restaurant	7
	Japanese Restaurant	4
	Chinese Restaurant	3
	Italian Restaurant	2
	Noodle House	2
	Soba Restaurant	2
	Tonkatsu Restaurant	2
	Bed & Breakfast	1
	Brazilian Restaurant	1
	Burger Joint	1
	Café	1
	Hotpot Restaurant	1
	Kushikatsu Restaurant	1
	Pizza Place	1
	Russian Restaurant	1
	Sushi Restaurant	1
	Thai Restaurant	1
	Unagi Restaurant	1
	Yakitori Restaurant	1
	Yoshoku Restaurant	1
Ueno	Ramen Restaurant	6
	Japanese Restaurant	5
	Chinese Restaurant	4
	Café	3
	Bed & Breakfast	2
	Italian Restaurant	2
	Soba Restaurant	2
	Sushi Restaurant	2
	Thai Restaurant	2
	Tonkatsu Restaurant	2
	Yoshoku Restaurant	2
	American Restaurant	1
	Coffee Shop	1
	Donburi Restaurant	1
	Halal Restaurant	1
	Pizza Place	1
	Restaurant	1
	Seafood Restaurant	1

	Tempura Restaurant	1
	Unagi Restaurant	1
	Vegetarian / Vegan Restaurant	1

----Akabane----

	venue	freq
0	Ramen Restaurant	0.17
1	Soba Restaurant	0.11
2	Coffee Shop	0.08

----Arakawa----

	venue	freq
0	Chinese Restaurant	0.29
1	Ramen Restaurant	0.21
2	Japanese Restaurant	0.07

----Hongo----

	venue	freq
0	Café	0.25
1	Mediterranean Restaurant	0.12
2	Japanese Restaurant	0.12

----Ikebukuro----

	venue	freq
0	Café	0.16
1	Japanese Restaurant	0.12
2	Sushi Restaurant	0.12

----Itabashi----

	venue	freq
0	Chinese Restaurant	0.67
1	Restaurant	0.33
2	American Restaurant	0.00

----Kinshicho----

	venue	freq
0	Ramen Restaurant	0.19
1	Coffee Shop	0.12
2	Thai Restaurant	0.06

----Koenji----

	venue	freq
0	Ramen Restaurant	0.18
1	Café	0.16
2	Italian Restaurant	0.12

----Meguro----

	venue	freq
0	Chinese Restaurant	0.3
1	Japanese Restaurant	0.2
2	Coffee Shop	0.2

----Nagatacho----

	venue	freq
0	Coffee Shop	0.27
1	Japanese Restaurant	0.20
2	Pizza Place	0.07

----Nakano----

	venue	freq
0	Ramen Restaurant	0.25
1	Japanese Restaurant	0.19
2	Café	0.12

----Nerima----

```
        venue  freq
0      Café  0.5
1 Chinese Restaurant  0.5
2 American Restaurant  0.0
```

----Nihonbashi----

```
        venue  freq
0      Coffee Shop  0.12
1 Japanese Restaurant  0.12
2 Yoshoku Restaurant  0.08
```

----Odaiba----

```
        venue  freq
0 Japanese Restaurant  0.27
1 Coffee Shop  0.18
2 Yakitori Restaurant  0.09
```

----Omori----

```
        venue  freq
0 Coffee Shop  0.14
1 Ramen Restaurant  0.09
2 Café  0.09
```

----Setagaya----

```
        venue  freq
0 Café  0.25
1 Ramen Restaurant  0.17
2 Szechuan Restaurant  0.08
```

----Shibuya----

```
        venue  freq
0 Café  0.38
1 Japanese Restaurant  0.10
2 Chinese Restaurant  0.10
```

----Shinagawa----

```
        venue  freq
0 Donburi Restaurant  0.29
1 Sushi Restaurant  0.14
2 Japanese Restaurant  0.14
```

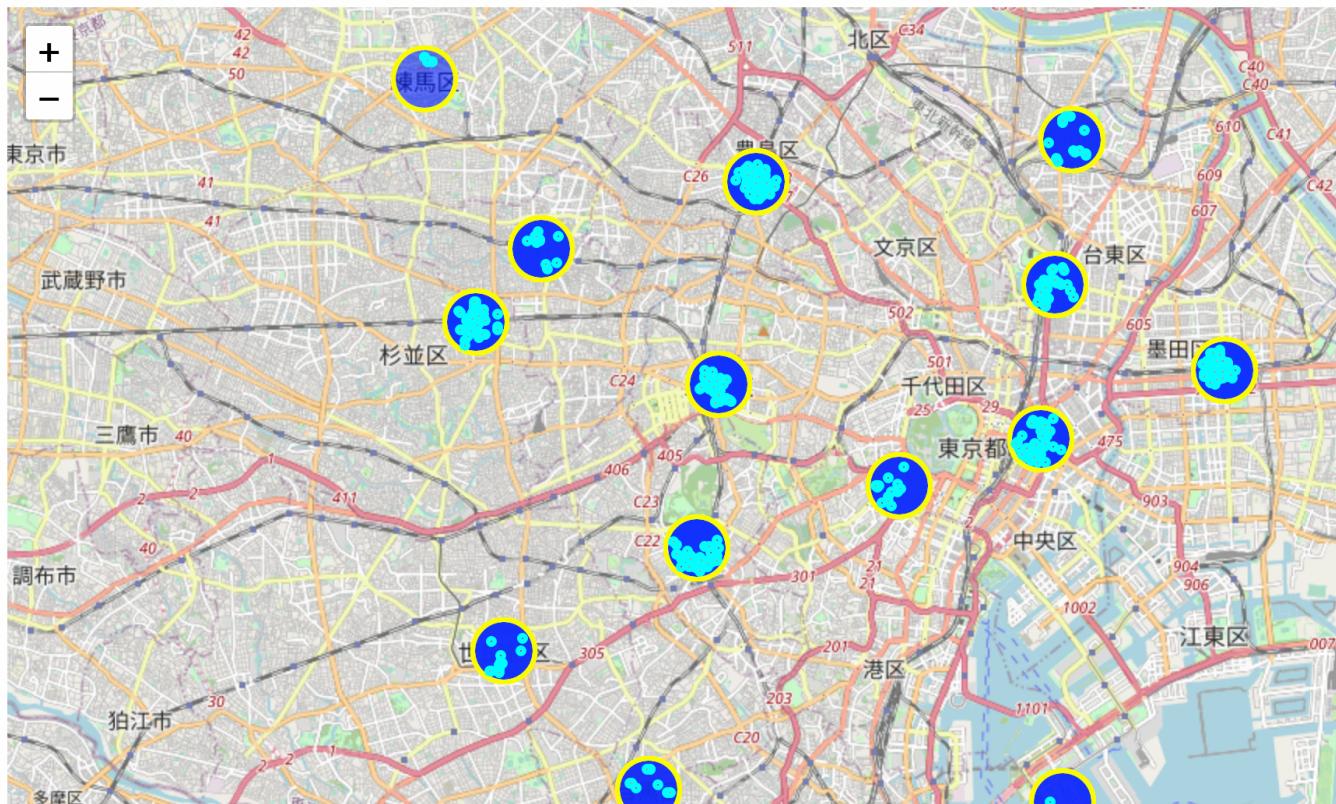
----Shinjuku----

```
        venue  freq
0 Ramen Restaurant  0.20
1 Japanese Restaurant  0.11
2 Chinese Restaurant  0.09
```

----Ueno----

```
        venue  freq
0 Ramen Restaurant  0.14
1 Japanese Restaurant  0.12
2 Chinese Restaurant  0.10
```

Folium Mapping Visualisation



Open a restaurant in following highlighted areas would face competitive environment, although the crowd of the restaurants indicates high volume of customers

Clustering Neighborhoods in Tokyo

Based on the similarities of venue categories, 22 neighborhood will be clustered by using K-Means. Clustered neighborhoods will be labelled from 0-4.

Pre-processing

	Neighborhood	American Restaurant	Asian Restaurant	Bangladeshi Restaurant	Bed & Breakfast	Bistro	Brazilian Restaurant	Breakfast Spot	Burger Joint	Café	...	Restau
0	Akabane	0.0	0.00	0.0	0.0	0.027778	0.0	0.0	0.000000	0.055556	...	0.0
1	Arakawa	0.0	0.00	0.0	0.0	0.000000	0.0	0.0	0.071429	0.071429	...	0.0
2	Hongo	0.0	0.00	0.0	0.0	0.000000	0.0	0.0	0.000000	0.250000	...	0.0
3	Ikebukuro	0.0	0.02	0.0	0.0	0.000000	0.0	0.0	0.000000	0.160000	...	0.0
4	Itabashi	0.0	0.00	0.0	0.0	0.000000	0.0	0.0	0.000000	0.000000	...	0.0

5 rows x 54 columns

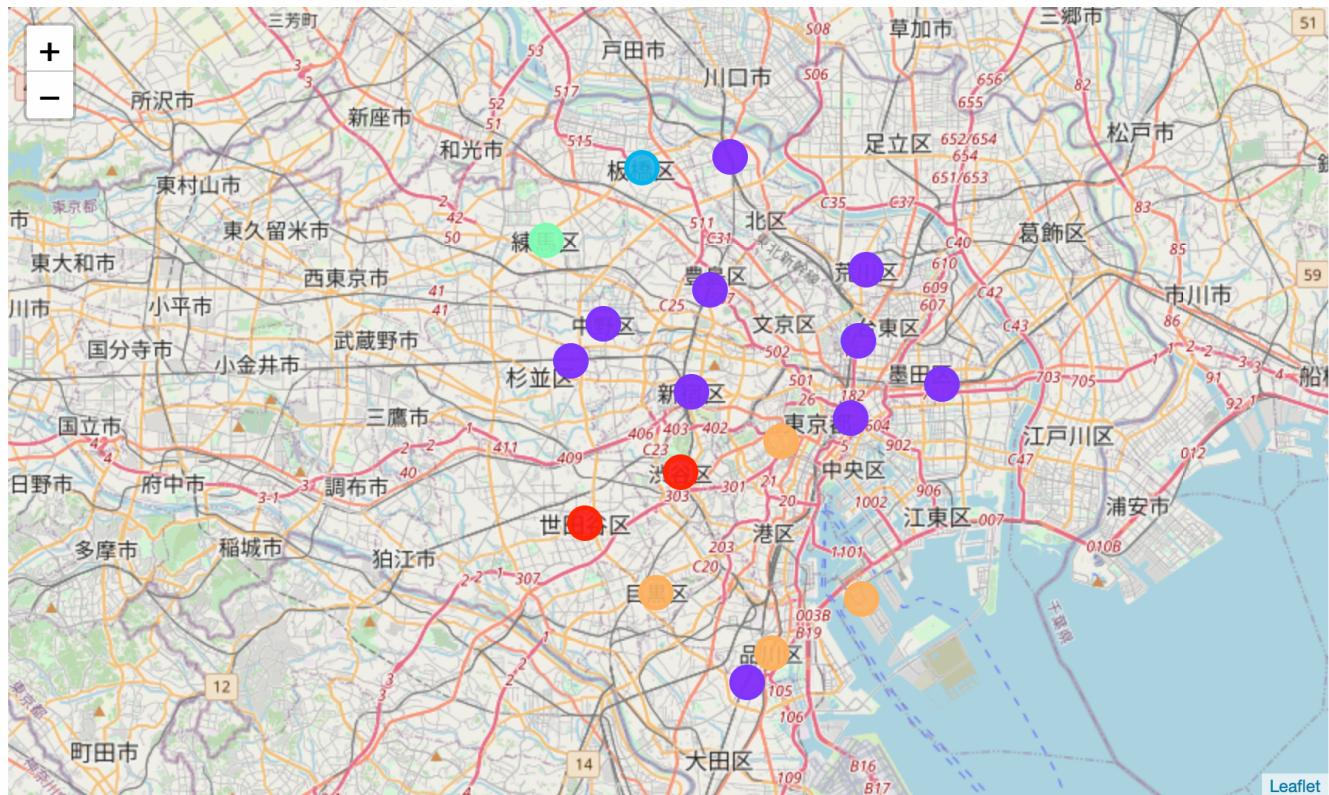
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Akabane	Ramen Restaurant	Soba Restaurant	Coffee Shop	Café	Japanese Restaurant	Indian Restaurant	Yakitori Restaurant	Dumpling Restaurant	Italian Restaurant	Kushikatsu Restaurant
1	Arakawa	Chinese Restaurant	Ramen Restaurant	Japanese Restaurant	Italian Restaurant	Noodle House	Burger Joint	Café	Coffee Shop	Donburi Restaurant	Fast Food Restaurant
2	Hongo	Café	Mediterranean Restaurant	Thai Restaurant	Pizza Place	Ramen Restaurant	Restaurant	Japanese Restaurant	Yoshoku Restaurant	Dumpling Restaurant	Italian Restaurant
3	Ikebukuro	Café	Japanese Restaurant	Sushi Restaurant	Yoshoku Restaurant	Ramen Restaurant	Coffee Shop	Chinese Restaurant	Dim Sum Restaurant	Italian Restaurant	Pet Café
4	Itabashi	Chinese Restaurant	Restaurant	Yoshoku Restaurant	Donburi Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hotpot Restaurant	Halal Restaurant

	Neighborhood	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	Akabane	Ramen Restaurant	Soba Restaurant	Coffee Shop	Café	Japanese Restaurant	Indian Restaurant	Yakitori Restaurant	Dumpling Restaurant	Italian Restaurant	Kushikatsu Restaurant
1	Arakawa	Chinese Restaurant	Ramen Restaurant	Japanese Restaurant	Italian Restaurant	Noodle House	Burger Joint	Café	Coffee Shop	Donburi Restaurant	Fast Food Restaurant
2	Hongo	Café	Mediterranean Restaurant	Thai Restaurant	Pizza Place	Ramen Restaurant	Restaurant	Japanese Restaurant	Yoshoku Restaurant	Dumpling Restaurant	Italian Restaurant
3	Ikebukuro	Café	Japanese Restaurant	Sushi Restaurant	Yoshoku Restaurant	Ramen Restaurant	Coffee Shop	Chinese Restaurant	Dim Sum Restaurant	Italian Restaurant	Pet Café
4	Itabashi	Chinese Restaurant	Restaurant	Yoshoku Restaurant	Donburi Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hotpot Restaurant	Halal Restaurant

	Name	Population	Density	Area	Major district	lat	long	Average Unit Price(t(JPY/sq.m))	China	Korea	...	1st Most Common Venue
0	Chiyoda	59441	5100	11.66	Nagatacho	35.675618	139.743469	2602846.0	1250.0	455.0	...	Coffee Shop
1	Chuo	147620	14460	10.21	Nihonbashi	35.684068	139.774503	2688537.0	3266.0	1401.0	...	Japanese Restaurant
2	Minato	248071	12180	20.37	Odaiba	35.619050	139.779364	2180846.0	3962.0	3461.0	...	Japanese Restaurant
3	Shinjuku	339211	18620	18.22	Shinjuku	35.693763	139.703632	940756.0	14153.0	10221.0	...	Ramen Restaurant
4	Bunkyo	223389	19790	11.29	Hongo	35.175376	137.013476	943847.0	4646.0	1658.0	...	Café

5 rows x 31 columns

Visualise all clusters



Display All Details of Each Cluster

Cluster: 0

	Major district	lat	long	Average Unit Price(t(JPY/sq.m)	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	...	1st Most Common Venue	2nd Most Common Venue
4	Hongo	35.175376	137.013476	943847.0	4646.0	1658.0	927.0	223.0	364.0	329.0	...	Café	Mediterranean Restaurant
11	Setagaya	35.646096	139.656270	665336.0	5835.0	4367.0	864.0	983.0	512.0	1706.0	...	Café	Ramen Restaurant
12	Shibuya	35.664596	139.698711	1360622.0	2013.0	1608.0	352.0	334.0	172.0	1383.0	...	Café	Japanese Restaurant

3 rows x 27 columns

Cluster: 1

	Major district	lat	long	Average Unit Price <small>(JPY/sq.m)</small>	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	...	1st Most Common Venue	2nd Most Common Venue
1	Nihonbashi	35.684068	139.774503	2688537.0	3266.0	1401.0	192.0	153.0	95.0	391.0	...	Japanese Restaurant	Chinese Restaurant
3	Shinjuku	35.693763	139.703632	940756.0	14153.0	10221.0	3484.0	747.0	3517.0	1033.0	...	Ramen Restaurant	Sushi Bar
5	Ueno	35.711759	139.777645	1025887.0	6489.0	3118.0	824.0	785.0	688.0	253.0	...	Ramen Restaurant	Japanese Restaurant
6	Kinshicho	35.696312	139.815043	607461.0	5874.0	1966.0	679.0	1349.0	320.0	183.0	...	Ramen Restaurant	Chinese Restaurant
10	Omori	35.588473	139.727933	557391.0	8467.0	3562.0	1494.0	2497.0	2287.0	621.0	...	Coffee Shop	Chinese Shop
13	Nakano	35.718123	139.664468	597965.0	6786.0	3365.0	1889.0	531.0	1750.0	496.0	...	Ramen Restaurant	Sushi Bar
14	Koenji	35.704942	139.649909	557008.0	5837.0	2854.0	1486.0	546.0	2226.0	750.0	...	Ramen Restaurant	Japanese Restaurant
15	Ikebukuro	35.730103	139.711884	732511.0	12955.0	2545.0	3609.0	531.0	3439.0	428.0	...	Café	Sushi Bar
16	Akabane	35.778139	139.720800	542465.0	10759.0	2432.0	2008.0	863.0	1342.0	247.0	...	Ramen Restaurant	Chinese Restaurant
17	Arakawa	35.737529	139.781310	472613.0	7284.0	5046.0	2075.0	528.0	1171.0	182.0	...	Chinese Restaurant	Japanese Restaurant

10 rows × 27 columns

Cluster: 2

	Major district	lat	long	Average Unit Price <small>(JPY/sq.m)</small>	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	...	1st Most Common Venue	2nd Most Common Venue
18	Itabashi	35.774143	139.681209	437376.0	14177.0	3218.0	1657.0	1545.0	1148.0	345.0	...	Chinese Restaurant	Japanese Restaurant

1 rows × 27 columns

Cluster: 3

	Major district	lat	long	Average Unit Price <small>(JPY/sq.m)</small>	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	...	1st Most Common Venue	2nd Most Common Venue
19	Nerima	35.74836	139.638735	400039.0	8177.0	4286.0	878.0	1099.0	690.0	527.0	...	Café	Chinese Restaurant

1 rows × 27 columns

Cluster: 4

	Major district	lat	long	Average Unit Price <small>t(JPY/sq.m)</small>	China	Korea	Vietnam	Philippines	Nepal	U.S.A.	...	1st Most Common Venue	2nd Com...
0	Nagatacho	35.675618	139.743469	2602846.0	1250.0	455.0	76.0	64.0	24.0	210.0	...	Coffee Shop	Japan Resta...
2	Odaiba	35.619050	139.779364	2180846.0	3962.0	3461.0	144.0	1027.0	107.0	3257.0	...	Japanese Restaurant	Coffee Shop
8	Shinagawa	35.599252	139.738910	766540.0	4317.0	2426.0	547.0	798.0	712.0	607.0	...	Donburi Restaurant	Soba Resta...
9	Meguro	35.621250	139.688014	863427.0	1836.0	1498.0	202.0	537.0	252.0	979.0	...	Chinese Restaurant	Coffee Shop

4 rows × 27 columns

Results

As stated at the beginning of this project, this report will answer the following questions by exploring the geographical and market data with the machine learning power deployed. * **Is the market saturated?**

Ramen restaurants are the most frequently distributed among all districts, which is followed by Cafe and Japanese restaurants, whereas Chinese restaurants are almost 50% less than the Ramen restaurants distribution frequency. It indicates: 1. Opening a Chinese restaurant is easier than that for a Ramen restaurant.

Because the market for Ramen restaurant are more likely to be saturated, which would bring fierce market competition. 2. Chinese restaurants might not as popular as Ramen and Japanese Restaurants. 3. Districts like Ikebukuro, Shinjuku, Ueno and Itabashi, that have higher Chinese residents proportions and relatively low land prices are most likely to be selected as the most strategical places to open a Chinese restaurant. Those districts are labelled as Cluster 2.

- **Who are the potential customers?**

Both Chinese residents and Non-Chinese residents are potential customers. The percentage of Chinese residents in Arakawa and Itabashi, where take the Chinese restaurant as their 1st Most Common Venue, are not the highest ones. This might point the Chinese restaurants are popular to not just Chinese Residents.

- **Who are the local competitors?**

Although Chinese restaurants are the 1st or 2nd popular restaurants for some districts, however there are at most 4 Chinese restaurants out of the popular restaurants of each districts. Chinese restaurants performed lower market occupancy than Ramen and Cafe, but they are much frequently distributed than all other kind of restaurants. Therefore, the main competitors of Chinese Restaurants are Japanese restaurants, especially Ramen restaurants.

According to my analysis, Cluster 2, where Chinese restaurants with low ranking of popularity, but with high Chinese residents proportions and fairly low land prices, are the most optimal locations to open a Chinese restaurant.

Discussion

Since the clustering was based only on the most common venues of each district, further discussions can be pivotting on other clustering such as DBSCAN. Some of the districts are not recorded in [TOKYO STATISTICAL YEARBOOK 2018](#). This kind of information might be found in Japanese websites. For the number of Chinese residents in Kasai and Tateshi could be discovered by who are interested and may understand Japanese.

Conclusion

This project offers a glimpse of how real life data-science projects look like. Despite the importance of traditional ways of doing marketing research, machine learning and data analyse techniques are largely facilitate the conveniences of solving business problems. My analysis were mainly based on the acquirable data from English websites, more interesting anslytics can be continuously conducted as more multilingual resources are surely available and will be discovered by interested readers.