Battle of neighborhood report: Start bubble tea business in Tokyo

1. Introduction

Tokyo is the capital of Japan and is one of the most attracting cities in the world. For many people travelling to Tokyo, they are impressive with Japanese food and restaurants. There was a saying in a Japanese commercial decades ago which is: '3pm is the time for tea and sweets.' Coffee, drinks, bread and desserts are always popular in this country.

Since 2018, bubble tea has brought a big boom in Tokyo. Although it is not a new drink, some new bubble tea shops have become trendy. People even line up for several hours to get a cup of delicious tapioca bubble tea. Due to the increasing influence of social media such as Twitter, Instagram, bubble tea shops can easily become popular over one night. To start the business in this category, we need to do some researches and analysis to help us start up better.

The business problem
The business problem I want to solve is to open a new bubble tea shop in Tokyo. Usually we would like to ask questions like: why? How? Where? This report will focus on why and where. Utilizing the data science techniques including data analysis, data visualization and machine learning, I would like to deliver the solutions for the following questions:
- Why should invest in bubble tea shop?
- Current bubble tea shop market analysis
- Where is the optimal location for the shop?

Target Audience
- Individuals interested in investment of bubble tea shop
- Bubble tea lovers
- Data scientist interested in data analysis and visualization

We will show you our analysis process step by step in this report.

2. Data preparation

In data preparation, I would like to collect the information for all bubble tea shops in Tokyo. Using Foursquares app, I will be able to find more venues information for different categories. For the first step, I would like to get the information about the wards and neighborhoods in Tokyo for further data collection.

2.1 Data for wards and neighborhoods

I find a sortable table on Wikipedia about all special wards and districts in Tokyo. This is a useful table and I would like to scrap the table by Python for further analysis. I used the powerful 'BeautifulSoup' package in python and get the raw data as following:

```
soup = BeautifulSoup(website_url,'lxml')
Tokyo_table = soup.find('table',{'class':'wikitable sortable'})
import pandas as pd
table_rows = Tokyo_table.find_all('tr')

data = []
for row in table_rows:
    data.append([t.text.strip() for t in row.find_all('td')])
#data formatting
df = pd.DataFrame(data)
df.drop([0],axis=0,inplace = True)
df.drop([24],axis=0,inplace = True)
df.drop([1,3,4,5,6],axis = 1, inplace = True)
df.columns = ['number','Ward','District']
df.head()
```

[177]:

	number	Ward	District
1	01	Chiyoda	Nagatachō, Kasumigaseki, Ōtemachi, Marunouchi,...
2	02	Chūō	Nihonbashi, Kayabachō, Ginza, Tsukiji, Hatchōb...
3	03	Minato	Odaiba, Shinbashi, Hamamatsuchō, Mita, Roppong...
4	04	Shinjuku	Shinjuku, Takadanobaba, Ōkubo, Kagurazaka, Ich...
5	05	Bunkyō	Hongō, Yayoi, Hakusan

I also would like to separate these districts into different rows. I split all the district columns into different columns and rearranged them into rows. The ward and district data finally look as the following. It simply contains two columns: ward and district included in different ward.

	District
Ward	
Chiyoda	Nagatachō
Chiyoda	Kasumigaseki
Chiyoda	Ōtemachi
Chiyoda	Marunouchi
Chiyoda	Akihabara
Chiyoda	Yūrakuchō
Chiyoda	Iidabashi
Chiyoda	Kanda
Chūō	Nihonbashi
Chūō	Kayabachō
Chūō	Ginza
Chūō	Tsukiji
Chūō	Hatchōbori
Chūō	Shinkawa

2.2 Longitude and latitude data

To generate the map or get the venues information, I also need to know the longitude and latitude information about these districts. I used 'Geopy' package and the data we got in section 2.1 to add corresponding latitude and longitude information to each district. Then we get the data:

	Ward	District	latitude	longtitude
0	Chiyoda	Nagatachō	35.675618	139.743469
1	Chiyoda	Kasumigaseki	35.672478	139.752388
2	Chiyoda	Ōtemachi	35.686794	139.766200
3	Chiyoda	Marunouchi	35.204882	136.841895
4	Chiyoda	Akihabara	35.699736	139.771250
5	Chiyoda	Yūrakuchō	35.674423	139.761674
6	Chiyoda	Iidabashi	35.701869	139.745381
7	Chiyoda	Kanda	29.855563	81.298341
8	Chūō	Nihonbashi	35.684058	139.774501
9	Chūō	Kayabachō	35.680229	139.779424

The information may not be complete that the districts covered in this project are only significant and major ones. In this report, I will explore 106 districts which we have the geographic data.

2.3 Generate the map

Utilizing the folium map package in Python, we can give a more user-friendly picture to show you the range of these districts we will research for this project. as shown in the following map:
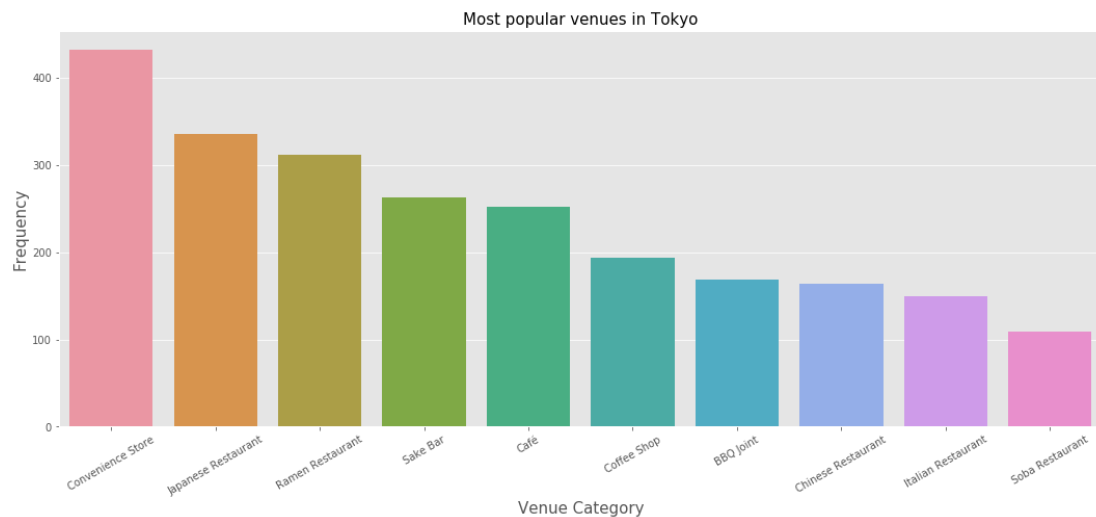
2.4 Foursquare app data

We can also get the venues data from Foursquare app. I collected 200 popular venues within one kilometer of each neighborhood.

3. Data analysis for beverage in Tokyo

3.1 Venue category analysis

People love food and drink in Tokyo. I calculated the top 10 popular venue categories in Tokyo. As shown in the chart below, 9 out of 10 are related to food and drink. Besides, Café and coffee shop related to beverage place as 5th and 6th.
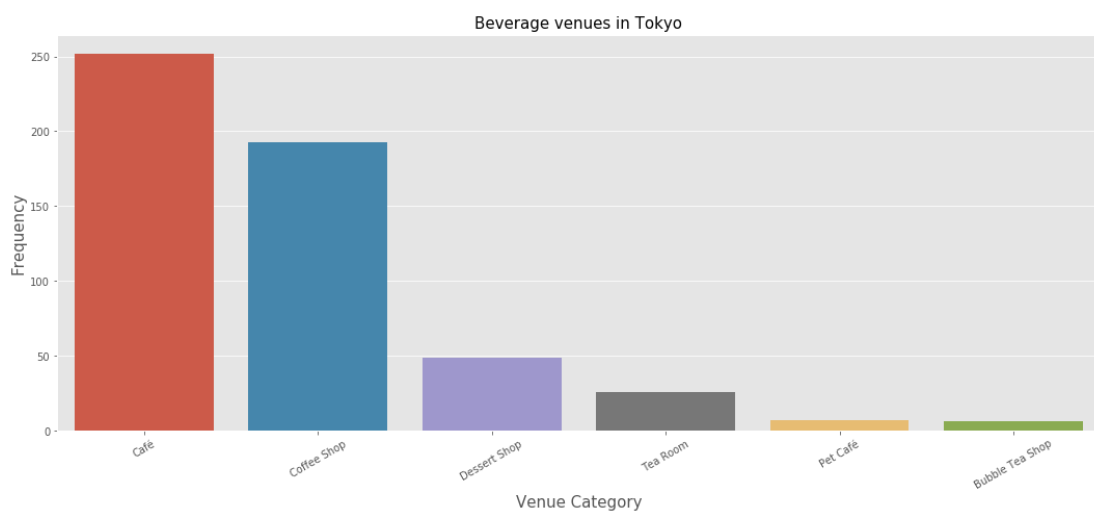

How about bubble tea? In our dataset, we only get 6 shops tagged as 'Bubble Tea shop' as shown below:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Akihabara	35.699736	139.771250	THE ALLEY	35.701925	139.772092	Bubble Tea Shop
Aomi	35.624779	139.781202	Chun Shui Tang (春水堂)	35.624824	139.779746	Bubble Tea Shop
Harajuku	35.668705	139.705336	Chun Shui Tang (春水堂)	35.668705	139.706977	Bubble Tea Shop
Harajuku	35.668705	139.705336	Gong cha (Gong cha 貢茶)	35.667431	139.705710	Bubble Tea Shop
Daikanyama	35.648157	139.703293	Chun Shui Tang (春水堂)	35.648445	139.702999	Bubble Tea Shop
Asagaya	35.704840	139.635472	Gong cha (Gong cha 貢茶)	35.704722	139.636499	Bubble Tea Shop

Surprisingly it is a small number compared to the popularity. I also checked yelp.com, there are 86 results for 'bubble tea' search. The reason why our results are few may due to:

1) Bubble tea shops are over 1 kilometer away from these neighborhood
2) Some shops sell bubble tea, but they have a different venue category.

To expand our results, I also checked the following words:" cafe"," dessert"," tea".

The total number of bubble tea shops are fewer than café and coffee shop. However, it seems to explain why people need to line up for hours to buy bubble tea. The demand is huge, but the supply is still low.
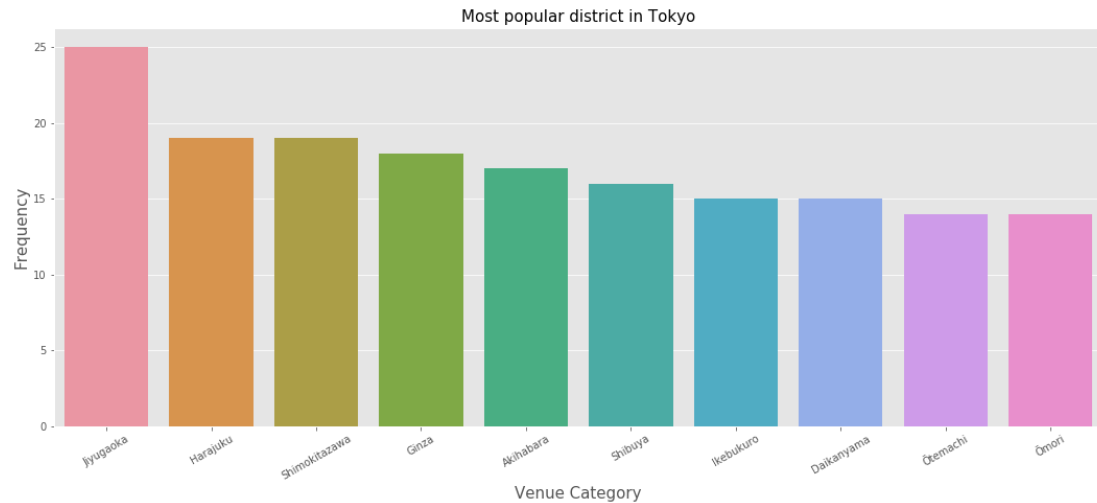
To conclude,

• Food and drink related venues are very popular in Tokyo.
• The number of specialized bubble tea shop in major neighborhood is low.
• Compared to café and coffee shop, there are still large gaps between demand and supply for bubble tea shop.
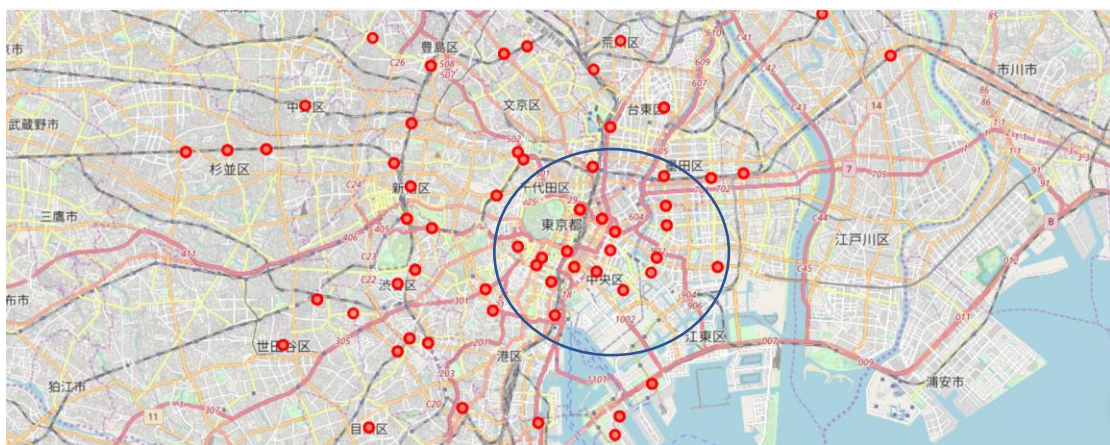
3.2 District distribution analysis

For this section, I would like to dig into the location a bit more. First, I find out the most popular top 10 districts for beverage in Tokyo as below:
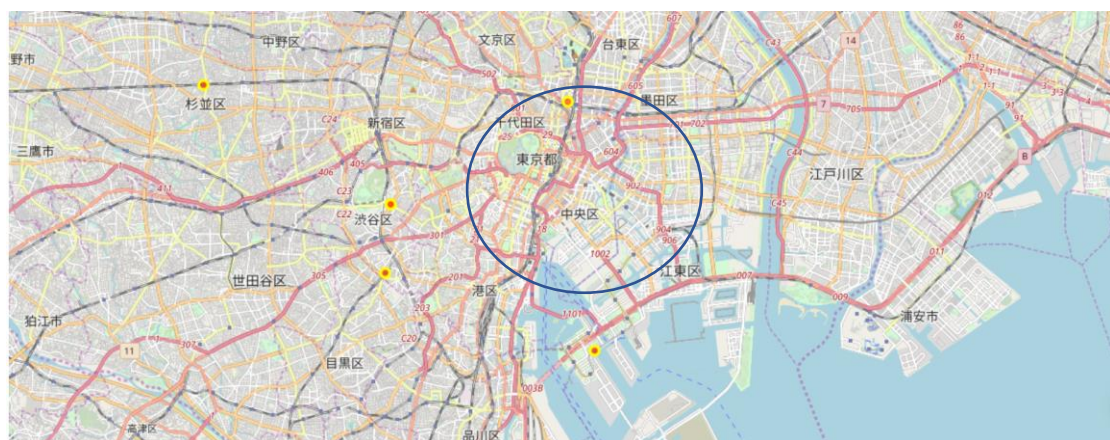
Most popular district in Tokyo

The top 3 popular districts are Jiyugaoka, Harajuku and Shimokitazawa. Jiyogaoka has more popular shops for beverage. Recall the 6 shops we have found, 2 of them are in Harajuku, which is the 2nd popular district. For the location, Jiyugaoka may be a good choice for a new bubble tea shop.

I also show these popular venues as a map above, most of the shops are distributed. Most of these shops are located in the middle of Tokyo. However, if we check the situation for bubble tea:
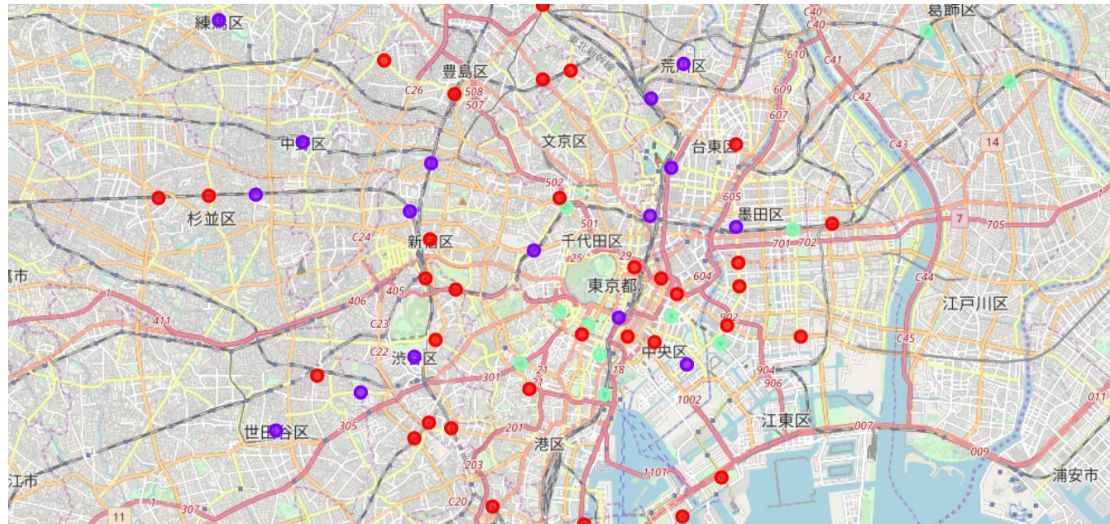

Only one shop found in the area circled.

To conclude, we find that the bubble tea shops are mostly absent from these popular areas,

like Jiyugaoka.

4. Cluster the neighborhood

Utilizing the K-means clustering, I classify 3 clusters based on the similarity of these venues in different districts. Purple dots are districts dominated by coffee shops, and read dots are districts dominated by café. Green dots are having mixed features, but 'tea' and 'desserts' seems to be a common feature for these districts. Green dots are relatively fewer than other colors, which can be a target market for our investment. Dessert and tea lovers usually have more interest in bubble teas.

Considering the convenience of transportation, Roppongi and Meguro are also good locations.

5. Conclusion and results

The reason why to invest in bubble tea shop is basically due to the gap between the demand and supply. Not a lot of specialized bubble tea shops cause the long line-up. There are large business opportunities in the bubble tea field. Of course, the technique to make a cup of delicious bubble tea is important, which is not shown in this report.

There are lots of popular café and coffee shops, but popular bubble tea shops are fewer. Compared to other categories under beverage, there is a huge space for the bubble tea business to grow in the future.

Otherwise, the bubble tea shops do not appear in most of these popular areas for beverage. However, these popular shops are all near railways or metros which people can easily get by public transportation.

As suggestions, I believe for our business problems raised at the very beginning:
- It is valuable to consider investment in bubble tea shop due to the large supply and low

competition situation.

- Beverage is popular, but bubble tea is still a small category based on the market analysis
- **Jiyugaoka** is a good choice for the intensity and popularity of beverage in this area. Furthermore, popular bubble tea shops are still absent here.
- Based on clustering analysis, **Roppongi and Meguro** have mixed features about beverage and prefer sweet desserts. They also have railway and subway nearby which can be good options to open new bubble tea shop.