

Battle of Neighborhoods–Tokyo

May 5, 2020

1 Introduction of Problem and Background

Today tourism is one of the most important parts of the economy. Tourists always want to travel to as many places as possible. On the other hand, there are too many amazing and exciting countries or cities to visit. Due to the limit of time or money, we have to make choices. Even after the travel destinations are determined, usually, there are still much more spots worth visiting than tourists could do actually.

Then the information about the destinations would be very helpful for travelers to make a good travel plan. For example, by clustering places in a city into several clusters, travelers can choose more dissimilar places to visit. This is also valuable for the travel agencies, since it would help to attract customers.

In this capstone project, I plan to choose Tokyo as the travel destination and choose the top 9 sights suggested by Google and use the resource of Foursquare to get the top venues around these sights and then make a travel plan based on the analysis about the data.

The structure of the rest part of the report is as follows. The second section, “Data Description”, explains the data to be used. The third section, “Methodology and Results” displays the methods used to analyze the data and the results. In the fourth section, “Discussion”, I discuss my observation based on the analysis in the third section and give my recommendation about the travel in Tokyo. The final section, “Conclusion” is used to conclude my report.

2 Data Description

Here I choose Tokyo (Japan) as the destination to visit. First, I search “Tokyo+touris+places” on Google. Then by clicking “top sights in Tokyo”, I get the Top 9 Sight in Tokyo (Fig.1). By scrapping the website, I got the coordinates about the 9 sights in Table 1. Next, by using the

Table 1: Top 9 Sights in Tokyo, Japan

ID	name	latitude	longitude
1	Tokyo Skytree	35.7101	139.811
2	Senso-ji	35.7148	139.797
3	Tokyo Tower	35.6586	139.745
4	Meiji Jingu	35.6764	139.699
5	Tokyo Disneyland	35.6329	139.88
6	Odaiba	35.6206	139.781
7	Ueno Park	35.7155	139.774
8	Imperial Palace	35.6852	139.753
9	Shinjuku Gyoen National Garden	35.6852	139.71

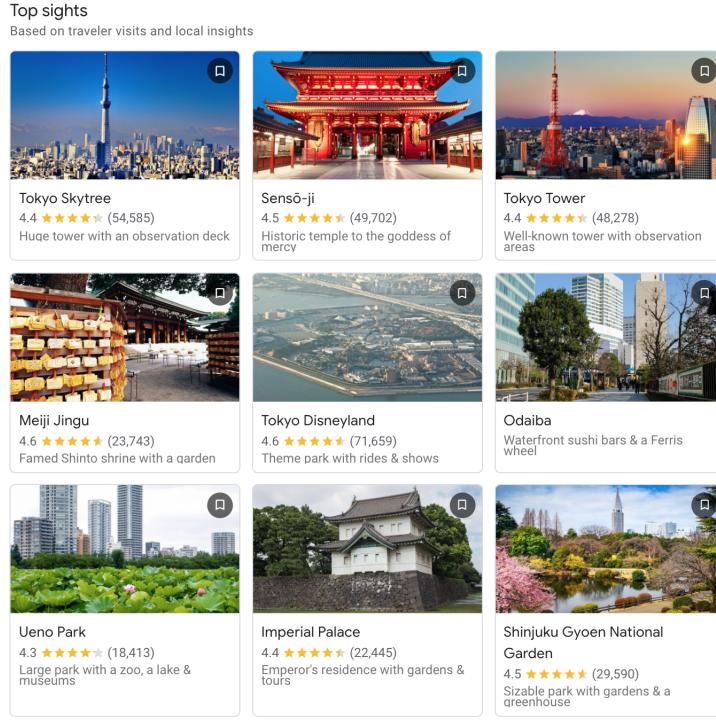


Fig. 1: Top 9 sights in Tokyo, Japan suggested by Google.

resource of Foursquare, I will get data about the venues around each sight. Then by getting the top 10 common venues in these places and clustering the top 9 sights into 5 clusters, a travel plan will be made.

3 Methodology and Results

The top 9 sights are visualized on the map by using Folium (Fig.2). Then by using the resource of the Foursquare, I get the top venues around each sight within 700 meters. The number of venues of one sight is limited to 100. Totally, 773 venues are obtained and the number of the unique venue categories is 179 (Fig. 3) .

By one-hot-vectorizing the ‘VenueCategory’ column over all of the venues and then by grouping rows by the name of the top 9 sights and taking the mean of the frequency of



Fig. 2: Map of Top 9 sights in Tokyo, Japan suggested by Google.

	name	sightLatitude	sightLongitude	Venue	VenueLatitude	VenueLongitude	VenueCategory
0	Tokyo Skytree	35.710063	139.8107	Tokyo Skytree (東京スカイツリー)	35.710054	139.810710	Monument / Landmark
1	Tokyo Skytree	35.710063	139.8107	Kitano Ace (北野エース)	35.710220	139.809799	Grocery Store
2	Tokyo Skytree	35.710063	139.8107	Sumida Aquarium (すみだ水族館)	35.710046	139.809639	Aquarium
3	Tokyo Skytree	35.710063	139.8107	Tokyo Skytree Tembo Galleria (東京スカイツリー天望回廊)	35.710132	139.810730	Scenic Lookout
4	Tokyo Skytree	35.710063	139.8107	Moomin House Cafe (ムーミンハウスカフェ)	35.710404	139.809349	Café

Fig. 3: Venues around the top 9 sights.

	name	ATM	American Restaurant	Aquarium	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Auditorium	BBQ Joint	Bakery	Bar	Bath House	Bed & Breakfast	Beer Bar	Beer Garden	Beijing Restaurant	bike Rental / Bike Share	Bike Shop	Bistro	Boat or Ferry	Bot Bookstore
0	Imperial Palace	0.00	0.000000	0.00	0.000000	0.066667	0.00	0.00	0.0000	0.000000	0.000000	0.000000	0.00	0.00	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.0000	
1	Meiji Jingu	0.00	0.000000	0.00	0.015152	0.000000	0.00	0.00	0.0000	0.030303	0.030303	0.030303	0.00	0.00	0.015152	0.00	0.00	0.015152	0.015152	0.000000	0.0000	
2	Odaiba	0.00	0.000000	0.00	0.000000	0.012500	0.00	0.00	0.0125	0.012500	0.000000	0.025000	0.00	0.00	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.0125	
3	Senso-ji	0.00	0.000000	0.00	0.000000	0.000000	0.01	0.00	0.0000	0.050000	0.000000	0.020000	0.01	0.01	0.000000	0.01	0.00	0.000000	0.000000	0.000000	0.0000	
4	Shinjuku Gyoen National Garden	0.01	0.000000	0.00	0.010000	0.000000	0.00	0.00	0.0000	0.030000	0.000000	0.030000	0.00	0.00	0.020000	0.00	0.01	0.000000	0.000000	0.010000	0.0000	
5	Tokyo Disneyland	0.00	0.020000	0.00	0.000000	0.000000	0.00	0.01	0.0000	0.000000	0.000000	0.010000	0.00	0.00	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.00	
6	Tokyo Skytree	0.00	0.010000	0.01	0.000000	0.000000	0.01	0.00	0.0000	0.040000	0.010000	0.010000	0.02	0.00	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.00	
7	Tokyo Tower	0.00	0.020619	0.00	0.000000	0.000000	0.00	0.00	0.0000	0.010309	0.030928	0.000000	0.00	0.00	0.000000	0.00	0.00	0.000000	0.000000	0.010309	0.0000	
8	Ueno Park	0.00	0.010000	0.00	0.010000	0.050000	0.00	0.00	0.0000	0.010000	0.000000	0.000000	0.00	0.01	0.010000	0.00	0.00	0.000000	0.000000	0.000000	0.01	

Fig. 4: DataFrame of frequencies of venue categories around the top 9 sights.

occurrence of each category, I get the dataframe about the frequencies of venue categories of the top 9 sights (Fig.4).

From the frequency dataframe, I get the 10 most common venues around the top 9 sights. Further more, in order to understand the similarities or dissimilarities about the venues around the top 9 sights, I cluster the 9 sights into 5 clusters by applying the k-means algorithm to the frequency dataframe. The result is displayed in Fig.5. This is also visualized in the map by Folium, as displayed in Fig.6.

4 Discussion

Based on the results in the last section, we can get a global vision about the venues around the top 9 sights. First, according to the 10 most common venues in Fig.5, we know that around Tokyo Skytree, there are convenience stores, cafes and scenic lookout. And you would find many Japanese restaurants around Senso-ji. If you like Chinese food, you should look for it around Tokyo Tower. Around both Meiji Jingu and Shinjuku Gyoen National Garden, there are many cafes. But, around Meiji Jingu, you can also find many Italian restaurants. On the other hand, around Shinjuku Gyoen National Garden, you can find many ramen restaurants.

Around Tokyo Disneyland, there are theme parks for you to visit. And Odaiba is a good place for shopping due to the top common coffee shops and plazas. Both Ueno Park and

	name	latitude	longitude	ClusterLabel	1st_MCV	2nd_MCV	3rd_MCV	4th_MCV	5th_MCV	6th_MCV	7th_MCV	8th_MCV	9th_MCV	10th_MCV
0	Tokyo Skytree	35.7101	139.811	1	Convenience Store	Café	Ramen Restaurant	BBQ Joint	Scenic Lookout	Japanese Restaurant	Coffee Shop	Grocery Store	Gift Shop	Chinese Restaurant
1	Senso-ji	35.7148	139.797	4	Japanese Restaurant	BBQ Joint	Soba Restaurant	Wagashi Place	Café	Coffee Shop	Dessert Shop	Unagi Restaurant	Yoshoku Restaurant	Hotel
2	Tokyo Tower	35.6586	139.745	0	Chinese Restaurant	Convenience Store	Park	Japanese Restaurant	Italian Restaurant	Bakery	Pizza Place	Ramen Restaurant	Steakhouse	Grocery Store
3	Meiji Jingu	35.6764	139.699	4	Café	Italian Restaurant	Convenience Store	Japanese Restaurant	Rock Club	Park	Japanese Curry Restaurant	Indian Restaurant	Udon Restaurant	BBQ Joint
4	Tokyo Disneyland	35.6329	139.88	4	Theme Park Ride / Attraction	Theme Park	Café	Gift Shop	Restaurant	Hotel	Snack Place	Japanese Restaurant	Lounge	Theme Restaurant
5	Odaiba	35.6206	139.781	2	Coffee Shop	Plaza	Convenience Store	Exhibit	Japanese Restaurant	Pier	Gift Shop	Bar	Shopping Mall	Discount Store
6	Ueno Park	35.7155	139.774	0	Zoo Exhibit	Sake Bar	History Museum	Art Museum	Wagashi Place	Coffee Shop	Chinese Restaurant	Historic Site	Japanese Restaurant	Museum
7	Imperial Palace	35.6852	139.753	0	Historic Site	Art Museum	Bridge	Palace	Library	Track	Plaza	Castle	Garden	Park
8	Shinjuku Gyoen National Garden	35.6852	139.71	3	Café	Ramen Restaurant	Japanese Restaurant	Garden	Coffee Shop	Bar	Burger Joint	Indian Restaurant	Theater	Chinese Restaurant

Fig. 5: 10 most common venues with cluster labels around the top 9 sights.

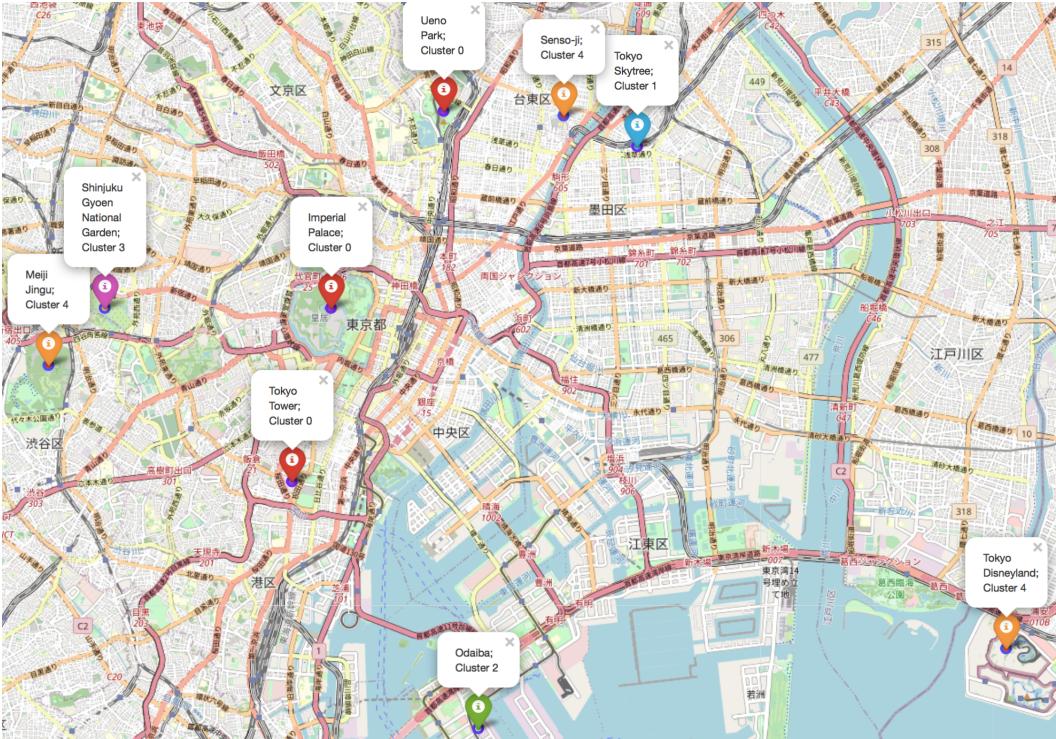


Fig. 6: Top 9 sights with cluster labels in Tokyo.

Imperial Palace are good places if you are interested in Japanese history and art. But, if, additionally, you like to taste some Japanese drink, you should go to Ueno Park, since around it there are Sake (Japanese alcohol) bars.

Of course, each of the top 9 sights is unique in its own. But, the result of the k-means clustering tells us that the venues around Senso-ji, Meiji Jingu and Tokyo Disneyland (Cluster 4) are very similar to each other. And this also holds among Tokyo tower, Ueno Park and Imperial Palace (Cluster 0). So, if the time is limited, you can choose one place to visit from each cluster.

Let me consider a tourist coming from other countries who is very interested in Japanese culture and plan to visit 5 places in Tokyo. Then I suggest a travel plan as follows. Firstly, it is Imperial Palace (Cluster 0), where there are historic sites, museums and palaces for him/her to experience the special Japanese history. Secondly, it is Shinjuku Gyoen National Garden (Cluster 3). Here, the tourist can visit the beautiful Japanese garden. Thirdly, it is Senso-ji (Cluster 4), where many Japanese food can be enjoyed (Soba Restaurant, Wagashi Place, Unagi Restaurant, etc.). Fourthly, it is Tokyo Skytree (Cluster 1), where from the Scenic Lookout, the tourist can enjoy the global vision of the beautiful Tokyo. Finally, it is Odaiba (Cluster 2) where is a good place for shopping!

5 Conclusion

In the above sections, starting from the top 9 sights suggested by Google, we get the venues around these sights by using the resource of Foursquare. Then we got the 10 most common venues around each sight and cluster these sights into 5 clusters by applying k-means to the frequencies of venue categories around these sights. The analysis gives us a global vision about the 9 sights. Based on it, some travel plans can be suggested. I think this will help tourists to have a better experience in Tokyo. This is also a good pattern for the travel agents to improve their services.