

Recommending Hotels for Olympics and Paralympic games in Tokyo 2021

Pali Masenkane

July 6, 2021

Introduction

A description of the problem and a discussion of the background

Following the covid19 pandemic outbreak, the Tokyo 2020 Olympic and Paralympics games were postponed by 12 months, and it has not been clear whether the games would continue this year, given that the pandemic is not yet over. Following this uncertainty, a decision has been made that the games will go ahead in Tokyo this year with Olympics expected to start on July 23, 2021, and end on August 8, 2021, while the Paralympics will start on the August 24, 2021 and end on September 5, 2021[1]. However, these games will take place under very different and strict restrictions as the virus continues to spread and mutate into different variants with slow access to the vaccines. In fact, a decision has been made that foreign spectators will not be allowed into Japan to watch the games. This means that the demand for accommodation and other related facilities has faded. Despite this, the organizing committee has decided that 10, 000 local spectators will be allowed per venue provided that this number is less 50% of the venue capacity. There will still be restrictions put in place in terms of movement restrictions and interactions. This means that the need for accommodation in Tokyo from among the more than 127 million people [2] staying in Japan will still be high and the high demand more specially for hotels closer to the games' venues especially the Tokyo Olympic stadium which is the main venue.

The problem, therefore, is to help local travelers with hotel information such as locations, reviews, and ratings so that they can book hotels that are conveniently located for their purpose.

A description of the data and how it will be used to solve the problem

For this exercise, Tokyo hotels listing data from Inside Airbnb[1] will be used to find and explore accommodation in Tokyo. These data list the hotel names and IDs, neighborhood, latitude and longitude, room types, price, reviews, and availability.

Source: [1] <http://www.insideairbnb.com/get-the-data-html>

Target Audience

The following groups of people maybe interested in this project:

1. All people who are currently in Japan and are intending to visit Tokyo for the Olympics and Paralympic games 2021

2. Tourists who are planning a visit to Tokyo
3. Aspiring Data Scientists who need to gain some experience in using FourSquare API in analysing location data

Data acquisition and cleaning

The data for this project, "Inside Airbnb" hotel listing data for Tokyo is used throughout this project. The dataset has been downloaded as a CSV file to a local computer and it was loaded from the local computer into Jupyter notebook as a platform that was used for this activity.

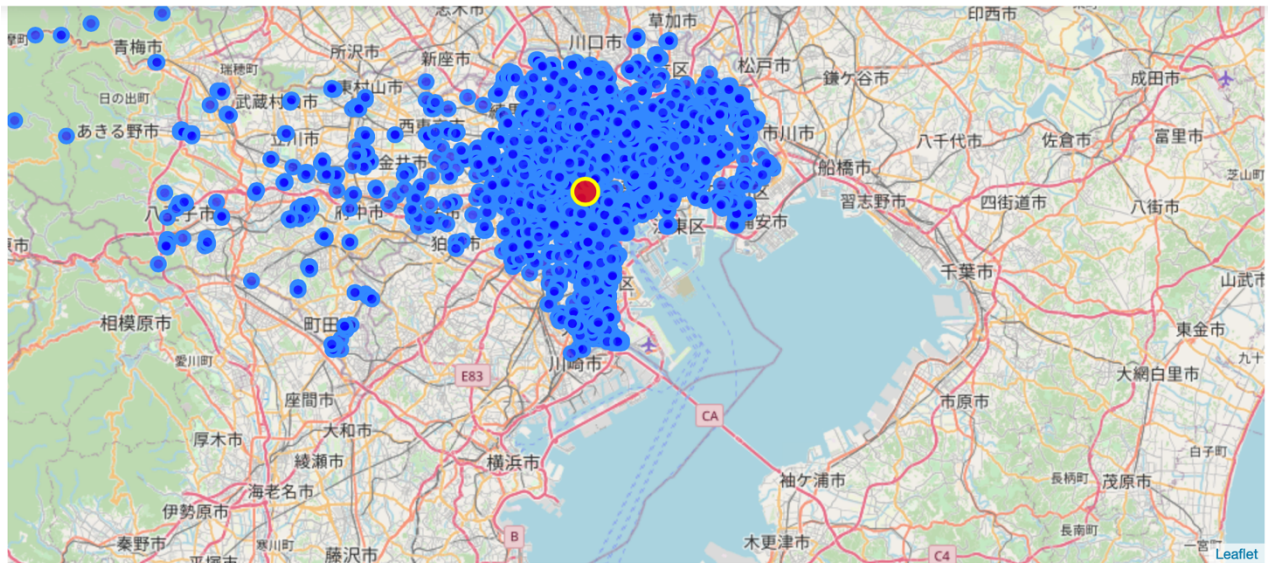
```
df_tokyo = pd.read_csv("/Users/palimasenkane/Downloads/listings.csv")
df_tokyo.head()
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	number_of_reviews
0	197677	Oshiage Holiday Apartment	964081	Yoshimi & Marek	NaN	Sumida Ku	35.71721	139.82596	Entire home/apt	11000		3
1	776070	Kero-kero house room 1	801494	Kei	NaN	Kita Ku	35.73844	139.76917	Private room	7677		3
2	899003	Classy room @Shinjuku, Takadanoba	4799233	Yu	NaN	Shinjuku Ku	35.71019	139.69755	Entire home/apt	5200		30
3	1016831	WOMAN ONLY LICENSED ! Cosy & Cat behnd Shibuya	5596383	Wakana	NaN	Setagaya Ku	35.65800	139.67134	Private room	4000		1
		private							Private			

Having studied the dataset size and the columns in it, some columns that were found unnecessary for this project were completely dropped from the dataset. Only those columns that would help us achieve the exercise were kept.

Other details of the dataset were relevant to the project and there was not much cleaning of the data that was done.

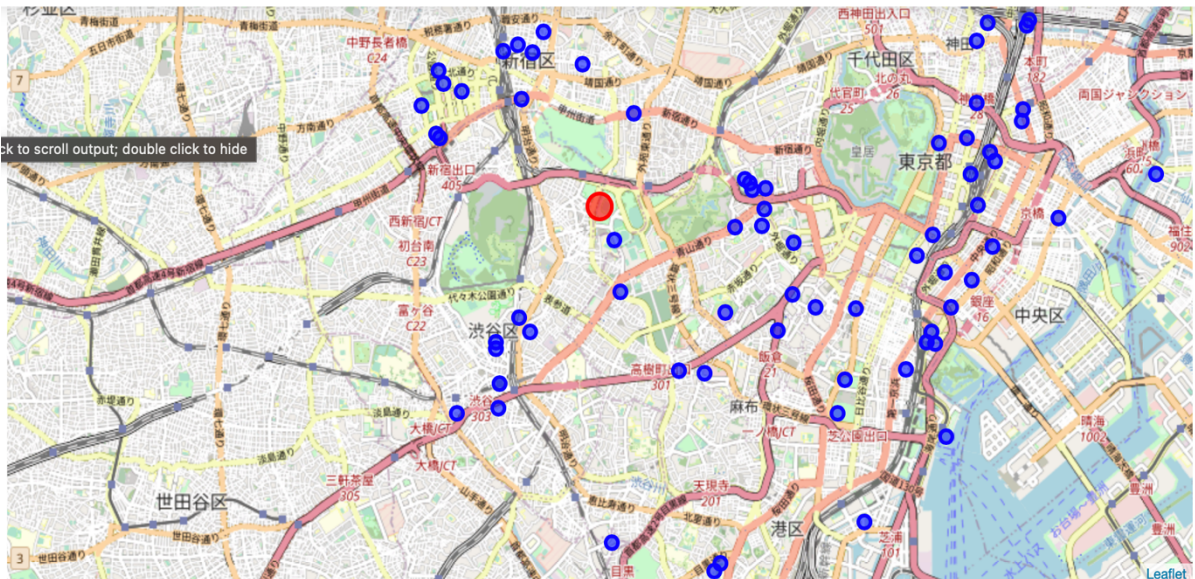
The Hotels in Tokyo were then visualized on the Tokyo Map centered around the Tokyo Olympic Stadium , which will be the main venue of the games thus a focal point.



The Methodology

A fourSqaure API has been used to find some hotels that are within a 10km distance of the Tokyo Olympic Stadium which will be the main venue for the games. Some hotels are also explored for find their ratings and tips from the users and to also locate some places of interests around the hotels.

The hotels are within the 10 km radius of the Tokyo Olympic Stadium were identified and visualized as shown in the figure below. The red circle indicates this stadium while the blue circles are these hotels.



Several of these hotels were also explored to find what could be some places of interests around them. The figure below shows some restaurants, convenient store, other hotels, gyms etc. and their exact physical addresses.

9	5731b405498e21762797d697	(アメリカンファーマシー)	'4bf58dd8d48988d10f951735', 'name': 'P...	1625480455	v-	False	紀尾井町1-3	東京ガーデンテラス 紀尾井町 1F	35.679200	139.736458
10	5bce982f031320002ccc1462	Freshness Burger (フレッシュネスバーガー)	'4bf58dd8d48988d16c941735', 'name': 'B...	1625480455	v-	False	紀尾井町1-3	東京ガーデンテラス 紀尾井町 2F	35.679387	139.736319
11	5731e420498eca7c09802ccf	Erick South (エリックサウス)	'54135bf5e4b08f3d2429dfde', 'name': 'S...	1625480455	v-	False	紀尾井町1-2	東京ガーデンテラス 紀尾井町 2F	35.679187	139.737138
12	5f3736891737e93ff59ec4e	イーク紀尾井町	'4bf58dd8d48988d177941735', 'name': 'D...	1625480455	v-	False	紀尾井町1-3	東京ガーデンテラス 紀尾井町 3F	35.679343	139.736966
13	5731d701498ed5592d7d4bc1	Starbucks	'4bf58dd8d48988d1e0931735', 'name': 'C...	1625480455	v-	False	紀尾井町1-2	東京ガーデンテラス 紀尾井町 2F	35.679223	139.737529
14	4bb9d1b77421a593be07c340	紀伊和歌山藩徳川家屋敷跡	'4deefb944765f83613cdba6e', 'name': 'H...	1625480455	v-	False	紀尾井町1-2	NaN	35.679822	139.737106
15	57315908498e672677d06b91	DEAN & DELUCA	'4bf58dd8d48988d16d941735', 'name': 'C...	1625480455	v-	False	紀尾井町1-2	東京ガーデンテラス 紀尾井町 1F	35.679209	139.736652
		New Otani Garden								

Furthermore, some hotels and these places of interest were also explored for ratings and tips and such rating information is very important in informing the customer decisions of renting a hotel.

Results and Discussion

1. We have successfully managed to use the Tokyo Hotel listing data and the foursquare api to pick hotels around Tokyo that are within 10 km of the Tokyo Olympic Stadium, the main venue for Tokyo2020 Olympic and Paralympic games, that will, among other host both the opening and closing ceremonies of the events.
2. We did not only list the hotels but we also managed to visualise them on the map centered around the Tokyo Olympic stadium and by simply putting the cursor on each blue spot, the name of the hotel is displayed both in English and Japanese making it easy for both Japanese and English speakers to easily see the name and search for other information from the internet.
3. Again, as shown from the results, we also managed to successfully explore some places of interests around certain hotels that were picked randomly. The knowledge of these places further make it easy for travellers to decide on which hotel to pick based on their interests. Not only are these places of interest listed, the distance from the hotel as well as the exact physical address are provided (in Japanese and English).
4. Only challenge for this project was the venue ratings and tips information. We were not able to display the rating information as all the randomly chosen venues returned "The venue has not been rated yet" message. The same happened with the "tips" information. This is a very important piece of information which can be very helpful in choosing the hotel as well.

Conclusion

The purpose of this project was to help travellers, particularly the people in Japan who will be travelling to Tokyo for the Olympics and paralympic games to easily locate hotels which are within 10km of the main venue of the games. This is important because the games will be held under very strict covid19 conditions which will limit movements and use of public transport. Thus, choosing a venue that is very close to areas of interest is critical. Since the purpose was to help travellers locate hotels and other places of interest, there was no much of data analysis, machine learning training, fitting and predicting of the models; no statistical inferences. Our approach was informed by our purpose, which is an important part of a data scientist to always be vigilant in choosing the tools and approaches on a given scenario or problem.

While we have been largely successful in reaching our main for this project, there is a huge room for improvement and with more and more practice things will get better.