Exploring for potential New Hotel venues in Tokyo Districts

1. Introduction

1.1. Background

This report is for those who are planning to start a new hotel in the city of Tokyo. It provides a suggestion on what would be the best venue to start a new hotel in a highly visited city with already many good hotels. Tokyo is one of the most popular tourist destinations in the world. Tokyo is one of the busiest cities in the world. From 2010 to 2017, nearly 13.8 Million tourists visited Tokyo. Tokyo has many attractions and would be a very good place to start a hotel.

1.2. Business Problem

This report focusses on the issue of where to open a new hotel in a city like Tokyo, once one has decided to go ahead. Let's imagine the Tata Subsidiary Tata Hotels willing to open a new luxury hotel, a first and foremost important decision will be the location for its new hotel.

1.3. Interest

On what basis can Tata decide its new hotel's location?

While selecting the place there are key points to consider like they need to check out like where the most well-visited venues of the city are?

If in case there are already other luxury hotels which have good ratings, will it be risky to open new one near these hotels? ¶

2. Data Preparation

2.1. Data Source - Scrapping Tokyo Districts Table from Wikipedia

I first make use of Districts of Tokyo page from Wiki to scrap the table to create a data-frame. For this, I used requests and Beautifulsoup4 library to create a data-frame containing name of the districts of Tokyo, Area, population and 11 districts. The webpage used for this project is - https://simple.wikipedia.org/wiki/Special_wards_of_Tokyo

2.2. Data Cleaning

I will use the Beautiful soup package to scrap the data from the HTML page to form the necessary table details, from it I will obtain the top ten districts, their Area and Population respectively.

These details will be used throughout this project for us to analyse and obtain the necessary results.

3. Exploratory Data Analysis

3.1. Gathering Details about the Target Districts

Once the Data is scarped and cleansed, I will focus on the gathering of the Details for the Target districts in Tokyo. This will be top eleven districts as per their Alphabetical order.

	District	Size km²	Population	Latitude	Longitude
0	Adachi[4]	11,830.68	629,392	35.7750	139.8044
1	Arakawa[5]	18,262.25	194,777	35.7359	139.7835
2	Bunkyō[6]	16,009.28	194,933	35.7079	139.7524
3	Chiyoda[7]	3,763.06	43,802	35.6940	139.7538
4	Chūō[8]	10,344.53	104,997	35.6706	139.7720
5	Edogawa[9]	13,264.86	661,386	35.7067	139.8682
6	Itabashi[10]	16,445.72	529,059	35.7512	139.7093
7	Katsushika[11]	12,286.62	428,066	35.7433	139.8472
8	Kita[12]	15,885.67	330,646	35.7529	139.7338
9	Kōtō[13]	10,963.24	436,337	35.6728	139.8175
10	Meguro[14]	18,217.55	267,798	35.6415	139.6981

Fig 1 – Table containing the 11 Districts in Tokyo with the necessary details.

3.2. Geographical view

Once I have got the necessary details of the data to be used in the project, the next step is to Fit the districts on a map as to get a Geographical view of the locations. The is very advantageous for this project as it will give a terrain view and a visual perspective of how the districts are placed.
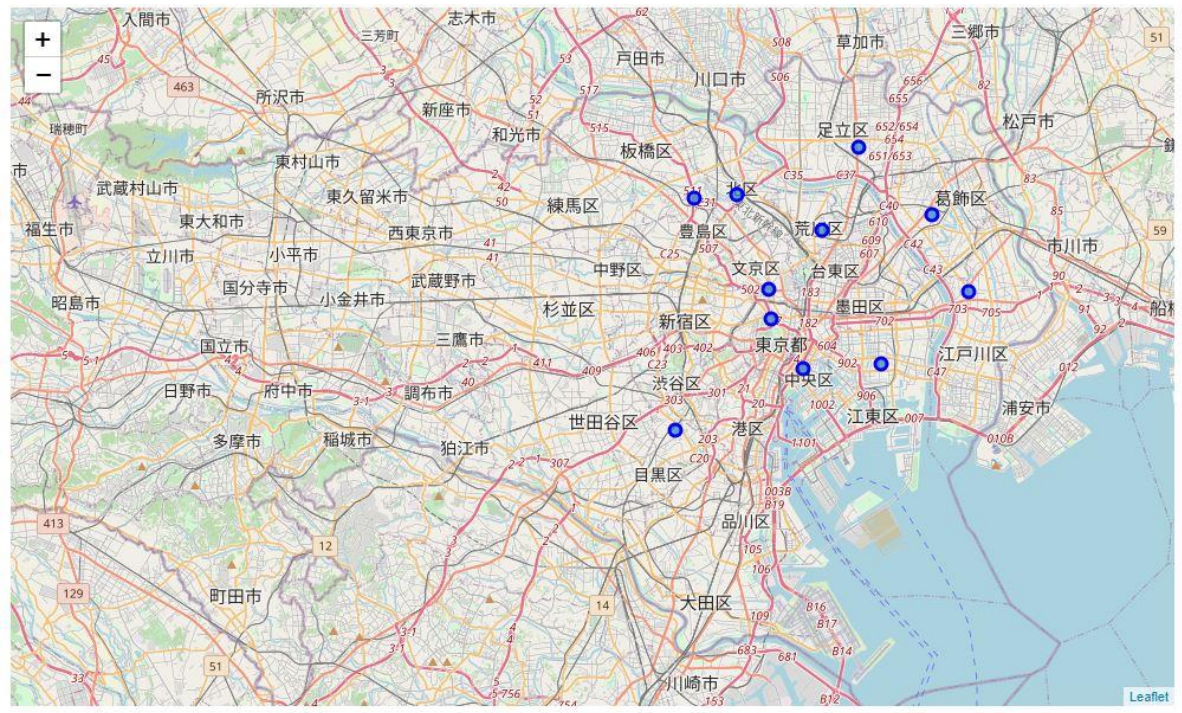

Fig 2 – Geographical View of the All the Districts in Tokyo.

4. Approach

4.1. Foursquare API

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

For this project, My Client ID and My Client Secret will be used to make a call to get the necessary Venue Latitude, Longitude and Category of Outlets in the Venue.

District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adachi[4]	31	31	31	31	31	31
Arakawa[5]	15	15	15	15	15	15
Bunkyō[6]	100	100	100	100	100	100
Chiyoda[7]	94	94	94	94	94	94
Chūō[8]	100	100	100	100	100	100
Edogawa[9]	28	28	28	28	28	28
Itabashi[10]	48	48	48	48	48	48
Katsushika[11]	28	28	28	28	28	28
Kita[12]	35	35	35	35	35	35
Kōtō[13]	85	85	85	85	85	85
Meguro[14]	95	95	95	95	95	95

Fig 3 – Table view of the Venue Latitude, Longitude and Number of Category Outlets in Each Districts.

4.3. Grouping Data

After getting the Venue Details through Foursquare, the next step I will group the Outlets in each Districts based on how common it is used. We will rank it for each District up to 10 in number.

	District	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	Adachi[4]	Convenience Store	Bus Stop	Park	Intersection	Grocery Store	Music Venue	Discount Store	Diner	Restaurant	Ramen Restaurant
1	Arakawa[5]	Convenience Store	Shipping Store	Tram Station	Grocery Store	Park	Sake Bar	Concert Hall	Chinese Restaurant	Café	Bus Stop
2	Bunkyō[6]	Baseball Stadium	Convenience Store	Japanese Restaurant	Ramen Restaurant	Coffee Shop	Theme Park Ride / Attraction	Concert Hall	BBQ Joint	Café	Discount Store
3	Chiyoda[7]	Café	Ramen Restaurant	Coffee Shop	Chinese Restaurant	Convenience Store	Japanese Curry Restaurant	Sake Bar	Italian Restaurant	Japanese Restaurant	French Restaurant
4	Chūō[8]	Japanese Restaurant	Coffee Shop	Soba Restaurant	Convenience Store	Sushi Restaurant	Ramen Restaurant	Bed & Breakfast	Chinese Restaurant	Café	Italian Restaurant
5	Edogawa[9]	Convenience Store	Grocery Store	Donburi Restaurant	Cultural Center	Bowling Alley	Indian Restaurant	Furniture / Home Store	Pizza Place	Electronics Store	Drugstore
6	Itabashi[10]	Convenience Store	Ramen Restaurant	Indian Restaurant	Grocery Store	Noodle House	Café	Chinese Restaurant	Shopping Mall	Restaurant	Japanese Restaurant
7	Katsushika[11]	Convenience Store	Intersection	Japanese Restaurant	Discount Store	Bus Stop	Supermarket	Sushi Restaurant	Pool Hall	Clothing Store	Drugstore
8	Kita[12]	Convenience Store	Ramen Restaurant	Intersection	Park	Japanese Restaurant	Theater	Fried Chicken Joint	Pizza Place	Pharmacy	Café
9	Kōtō[13]	Convenience Store	Japanese Restaurant	Coffee Shop	Ramen Restaurant	Chinese Restaurant	Sake Bar	Supermarket	Sushi Restaurant	Donburi Restaurant	Park
10	Meguro[14]	Café	Japanese Restaurant	French Restaurant	Yakitori Restaurant	BBQ Joint	Ramen Restaurant	Italian Restaurant	Coffee Shop	Chinese Restaurant	Nabe Restaurant

Fig 4 – Table view of the Data Grouped based on how common it is for the District

4.2. Clustering Approach

To compare the similarities of the Districts, i decided to explore neighbourhoods, segment them, and group them into clusters to find similar neighbourhoods in a big city like Tokyo. To be able to do that, i need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

I will import the necessary the Clustering package – Sklearn and KMeans
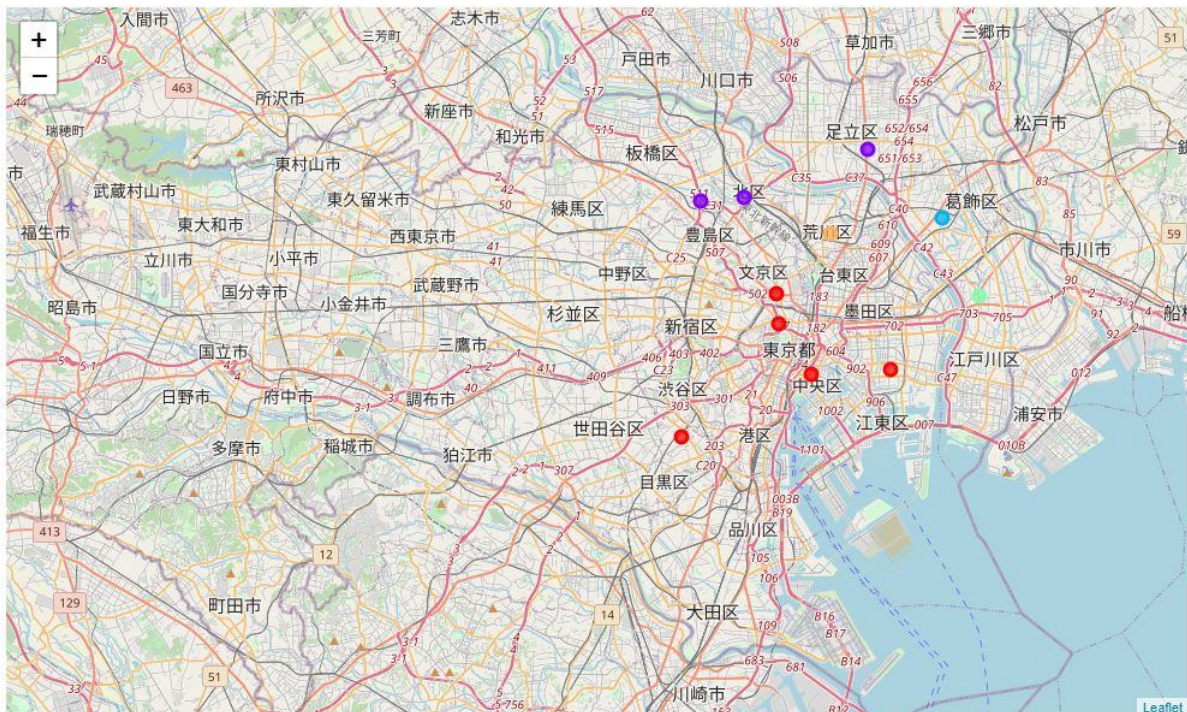
No of Clusters used – 5

Fig 5 – Map view of the Districts grouped into Clusters

4.3 Clusters

Cluster 0 – Restaurants

The first cluster will concentrate on the majority of the Restaurants situated in 5 Districts – Bunkyo, Chiyoda, Chuo, Koto and Meguro

District	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	9th Most Common Venue	10th M Comm Ve
Bunkyō[6]	35.7079	139.7524	0	Baseball Stadium	Convenience Store	Japanese Restaurant	Ramen Restaurant	Coffee Shop	Theme Park Ride / Attraction	Concert Hall	BBQ Joint	Café	Disco S
Chiyoda[7]	35.6940	139.7538	0	Café	Ramen Restaurant	Coffee Shop	Chinese Restaurant	Convenience Store	Japanese Curry Restaurant	Sake Bar	Italian Restaurant	Japanese Restaurant	Fre Restau
Chūō[8]	35.6706	139.7720	0	Japanese Restaurant	Coffee Shop	Soba Restaurant	Convenience Store	Sushi Restaurant	Ramen Restaurant	Bed & Breakfast	Chinese Restaurant	Café	Ita Restau
Kōtō[13]	35.6728	139.8175	0	Convenience Store	Japanese Restaurant	Coffee Shop	Ramen Restaurant	Chinese Restaurant	Sake Bar	Supermarket	Sushi Restaurant	Donburi Restaurant	F
Meguro[14]	35.6415	139.6981	0	Café	Japanese Restaurant	French Restaurant	Yakitori Restaurant	BBQ Joint	Ramen Restaurant	Italian Restaurant	Coffee Shop	Chinese Restaurant	N Restau

Fig 6 – Cluster 0 - Restaurants

Cluster 1 – Restaurants & Stores

The second cluster will concentrate on the Restaurants and Stores situated in 3 Districts – Adachi, Itabashi and Kita

	District	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	9th Most Common Venue	10th Most Common Venue
0	Adachi[4]	35.7750	139.8044	1	Convenience Store	Bus Stop	Park	Intersection	Grocery Store	Music Venue	Discount Store	Diner	Restaurant	Ramen Restaurant
6	Itabashi[10]	35.7512	139.7093	1	Convenience Store	Ramen Restaurant	Indian Restaurant	Grocery Store	Noodle House	Café	Chinese Restaurant	Shopping Mall	Restaurant	Japanese Restaurant
8	Kita[12]	35.7529	139.7338	1	Convenience Store	Ramen Restaurant	Intersection	Park	Japanese Restaurant	Theater	Fried Chicken Joint	Pizza Place	Pharmacy	Café

Fig 7 – Cluster 1 – Restaurants and Stores

Cluster 2 – Restaurants, Stores & Public Places

The third cluster will concentrate on the Restaurants, Stores and Public Places situated in the district of Katsushika.

	District	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	9th Most Common Venue	10th Most Common Venue
7	Katsushika[11]	35.7433	139.8472	2	Convenience Store	Intersection	Japanese Restaurant	Discount Store	Bus Stop	Supermarket	Sushi Restaurant	Pool Hall	Clothing Store	Drugstore

Fig 8 – Cluster 2 - Restaurants, Stores & Public Places

Cluster 3 - Restaurants & Stores

The fourth cluster will concentrate on the Restaurants and Stores situated in the district of Edogawa.

	District	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	9th Most Common Venue	10th Most Common Venue
5	Edogawa[9]	35.7067	139.8682	3	Convenience Store	Grocery Store	Donburi Restaurant	Cultural Center	Bowling Alley	Indian Restaurant	Furniture / Home Store	Pizza Place	Electronics Store	Drugstore

Fig 9 – Cluster 3 – Restaurants & Stores

Cluster 4 – Stores, Park and Train Station

The fifth cluster will concentrate on the Stores, Parks and Train Stations situated in the district of Arakawa.

	District	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	9th Most Common Venue	10th Most Common Venue
1	Arakawa[5]	35.7359	139.7835	4	Convenience Store	Shipping Store	Tram Station	Grocery Store	Park	Sake Bar	Concert Hall	Chinese Restaurant	Café	Bus Stop

Fig 10 – Cluster 4 – Stores, Park and Train Station

5. Results

The following are the highlights of the 5 clusters above:

Restaurants are clearly located only in Cluster 0 (Eastern district of Tokyo), which makes the choice of the final location very easy, in case Momentum does want to reduce risks. As for restaurants and stores for buying are very popular also in the Northern District of Tokyo in Cluster 1. Especially in Adachi district.

For Restaurants associated with Public Places, Cluster 2 has many places. Cluster 3 is similar to Cluster 2 which has many restaurants and stores together but here the shops are associated with Home items.

Cluster 4 Has minimal restaurants but has a few stores and has a bus stop and tram station.

6. Discussion

It is noticeable that Cluster 0 is the most viable clusters to build a new luxury hotel with guarantees. The proximity to a big number of Restaurants (lunch and dinner venues for guests), Coffee shops This cluster has five districts (Bunkyo, Chiyoda, Chuo, Koto, Meguro) with minimal Hotels.

7. Conclusion

In conclusion, this project would have had better results if there were more available data in terms of actual land pricing data within the area.

However, based on the available data, my advice to Tata Hotels would be to focus on only on Cluster 0 and within that in Bunkyo and Chiyoda Districts based on its Geographical Advantage when investing on a new luxury hotel.

8.Reference

1. https://simple.wikipedia.org/wiki/Special_wards_of_Tokyo - For Districts Details in Tokyo.

2.https://stats.stackexchange.com/questions/180883/k-means-clustering-not-working-as-expcected - Stakexchange - For K-Mean Reference.

3.https://www.codegrepper.com/code-examples/whatever/folium+map+python – Grepper – For Folium Maps Reference.

4. For project reference - "The Battle of Neighbourhoods" – Rdelhot – GitHub.