The Battle of Neighborhoods: Japan Hostels

1. Introduction

1.1 Background

As someone who loves travelling, finding ways to save money and travel low-cost is essential. There's little to save while booking flights and I'd rather pay more to traveling comfortably on a 16Hr flight than try to save and wear myself down even before the trip begins! But we can always save a few bucks on accommodation.

A hostel is a low-budget accommodation shared among travellers. Staying at hostels is an extremely affordable option for travellers. Not only that, It also gives you a chance to experience which you would never at a hotel. I've found that experiencing hostel life one of the best ways to travel the world on a budget and here's why:

1. You'll save a lot of money trust me.
2. Meet other like-minded travellers and make new friends around the world!
3. Most hostels organize lot of social activates

1.2 Business Understanding/Problem Description

Hostel accommodation is a booming industry with the latest rise in tourism worldwide, especially in Asia. Over 20% of the total tourists worldwide are aged between 19-25. And around 80% of them prefer to spend less on accommodation by opting for hostels. This means the demand for hostels is only going to grow further and more people would want to get their legs in the game.

- How should a new business person decide where to open a hostel?
- What factors should he look at before investing?
- Which neighborhood venues affect a user's rating for location of hostel?

At the same time, it is difficult for a travellers, especially first-timers, to select a hostel from among many options. Hostel reviews are subjective and differ from person-to-person and one cannot solely depend on them to make a decision. It is especially important to consider other aspects like price and neighborhood, which can greatly influence one's experience of the city/country. I will try to answer the following questions

- How does price vary with location?
- Where are the *value for money* hostels located?
- How does proximity to transportation affect hostel rating?
- Which hostels are most secure and where are they located?
- Suggest similar hostel but which cheaper price.

Tourism in Japan in on a rise. It is expected that the number of foreign tourists coming to Japan will be increasing till 2020 when Olympic will be held in Tokyo. Hence, for this project, we will be looking at hostels in Japan, in particular, Tokyo.

1.3 Target Audience

This project intends to serve two groups of audience:

1. **Travellers:** Help them make an informed decision while choosing a hostel by providing an in-depth analysis of hostels and their neighborhood.
2. **Business Person:** Provide useful information and models which can help them where to open their first/next hostel.

2. Data

I analysed in this project : **Tokyo**. Following are the datasets used in the project:

1. Japan Hostel Dataset
2. Hostel Neighborhood
3. Tokyo Land Price

2.1 Japan Hostel Dataset

The original dataset on Kaggle has the following columns:

- hostel.name: Hostel Name
- City: City name where hostel is located in
- price.from: Minimum Price for 1 night stay
- Distance: Distance from city center (km)
- summary.score: Summary score of ratings
- rating.band: Rating band
- atmosphere: Rating score of atmosphere
- cleanliness: Rating score of cleanliness
- facilities: Rating score of falicities
- location: Rating score of location
- security: Rating score of security
- staff: Rating score of staff
- valueformoney: Rating score of value for money
- lon: Longitude
- lat: Latitude

Below is a snapshot of the dataset:

Name	City	StartPrice	DistanceFromCityCentre	OverallScore	RatingCategory	Atmosphere	Cleanliness	Facilities	Location	Security	Staff
"Bike & Bed" CharinCo Hostel	Osaka	3300	2.9	9.2	Superb	8.9	9.4	9.3	8.9	9.0	9.4
&And Hostel Akihabara	Tokyo	3600	7.8	8.7	Fabulous	8.0	7.0	9.0	8.0	10.0	10.0
&And Hostel Ueno	Tokyo	2600	8.7	7.4	Very Good	8.0	7.5	7.5	7.5	7.0	8.0
&And Hostel-Asakusa North-	Tokyo	1500	10.5	9.4	Superb	9.5	9.5	9.0	9.0	9.5	10.0
1night1980hostel Tokyo	Tokyo	2100	9.4	7.0	Very Good	5.5	8.0	6.0	6.0	8.5	8.5

2.2 Hostel Neighborhood:

This dataset contains all the neighborhoods or venues within 500m radius of a Hostel. It has the following columns:

- HostelName: Name of the hostel
- VenueName: Name of the venue
- Category: It is the primary category of the venue, for eg. Café, Train Station, Restaurant.
- VenueLatitude, VenueLongitude: Coordinates of the venue.

Below is a snapshot of the dataset:

	HostelName	VenueName	Category	VenueLatitude	VenueLongitude
0	"Bike & Bed" CharinCo Hostel	Lawson (ローソン S OSL谷町四丁目駅北店)	Convenience Store	34.683282	135.517281
1	"Bike & Bed" CharinCo Hostel	モナコ	Café	34.683210	135.517309
2	"Bike & Bed" CharinCo Hostel	7-Eleven (セブンイレブン 大阪谷町3丁目店)	Convenience Store	34.683509	135.517793
3	"Bike & Bed" CharinCo Hostel	McDonald's (マクドナルド 地下鉄谷町四丁目店)	Fast Food Restaurant	34.683109	135.517443
4	"Bike & Bed" CharinCo Hostel	City Plaza Osaka (シティプラザ大阪)	Hotel	34.684020	135.510206

2.3 Tokyo Land Price:

This dataset contains the locality name and the average price of land per square meter. Below is a snapshot:

	Neighborhood	PricePerSqMeter
0	Chiyoda-Ku	1890610
1	Chuo-Ku	3178147
2	Minato-Ku	2339310
3	Shinjuku-Ku	930080
4	Bunkyo-Ku	966787

3. Methodology

3.1 Data Collection:

- The Japan Hostels dataset is freely available on Kaggle and was built by scraping Hostelworld.com website.
- We used Foursquare API to get the venues around the hostel.
- We scraped https://utinokati.com to get land prices of various neighborhoods in Tokyo.

3.2 Analytic Approach:

I took two approaches in the project.

Firstly, I used *exploratory data analysis(EDA)* to uncover hidden properties of data and provide useful insights to the reader, both future traveller and investor. I used the list of hostels from *Hostel dataset* and use *Foursquare API* to get venues around the Hostel. I will then use EDA to explore the neighborhood and how it affects the price of the hostel. I will also use the combined dataset to cluster similar hostels as per pricing and neighborhood.

Secondly, I used *prescriptive analytics* to help a business person decide a location for new hostel. I will use *clustering* (KMeans). I combined the above data with the land price for the area in which the Hostel is situated and them develop clustering models to predict where a new hostel should be opened.

4. Analysis

4.1 Exploratory Data Analysis

4.2 Clustering

5. Result

6. Discussion

7. Conclusion