Hostels of Japan

Introduction

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

Data

- I analysed in this project : Tokyo. Following are the datasets used in the project:
  - Japan Hostel Dataset
  - Hostel Neighbourhood
  - Tokyo Land Price

Methodology

- 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 neighbourhood and how it affects the price of the hostel. I will also use the combined dataset to cluster similar hostels as per pricing and neighbourhood.

  - Secondly, I used prescriptive analytics to help a business person decide a location for new hostel. I will use clustering(K-Means). 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.

Results

- We got a glimpse of the hostel scene in Tokyo and were able to find out some interesting insights which might be useful to travellers as well as people with business interests. Let's summarize our findings:
- Most hostels are located in Taito-ku and Chuo-ku.
- Sumida-ku seems to be an interesting locality since it is close to Taito-ku and 43% cheaper than Taito-ku.
- The starting price of hostels does not vary much depending on its distance from the city centre.
- Most of the hostels rated high for their security are in Katsushika, Kita, Meguro, Shibuya and Shinagawa
- Proximity to a mode of transportation or a historic site positively affects the hostel rating.
- Hostels rated highly for being value for money are comparatively cheaper and are located away from the city centre

Conclusion

- In the above study, we explored and analysed various aspects of the hostels scene in Tokyo, Japan using data science. We used an existing dataset and combined it with data collected from Foursquare API as well as data scraped from a website. We performed EDA and clustering on these datasets in our pursuit of solutions. We were able to find satisfactory answers to the questions we posed before the study.
-
- The study is based on limited data, but it is nevertheless a significant step in shedding light on the hostel scene in Tokyo. This study can be repeated easily for other cities of the world.