

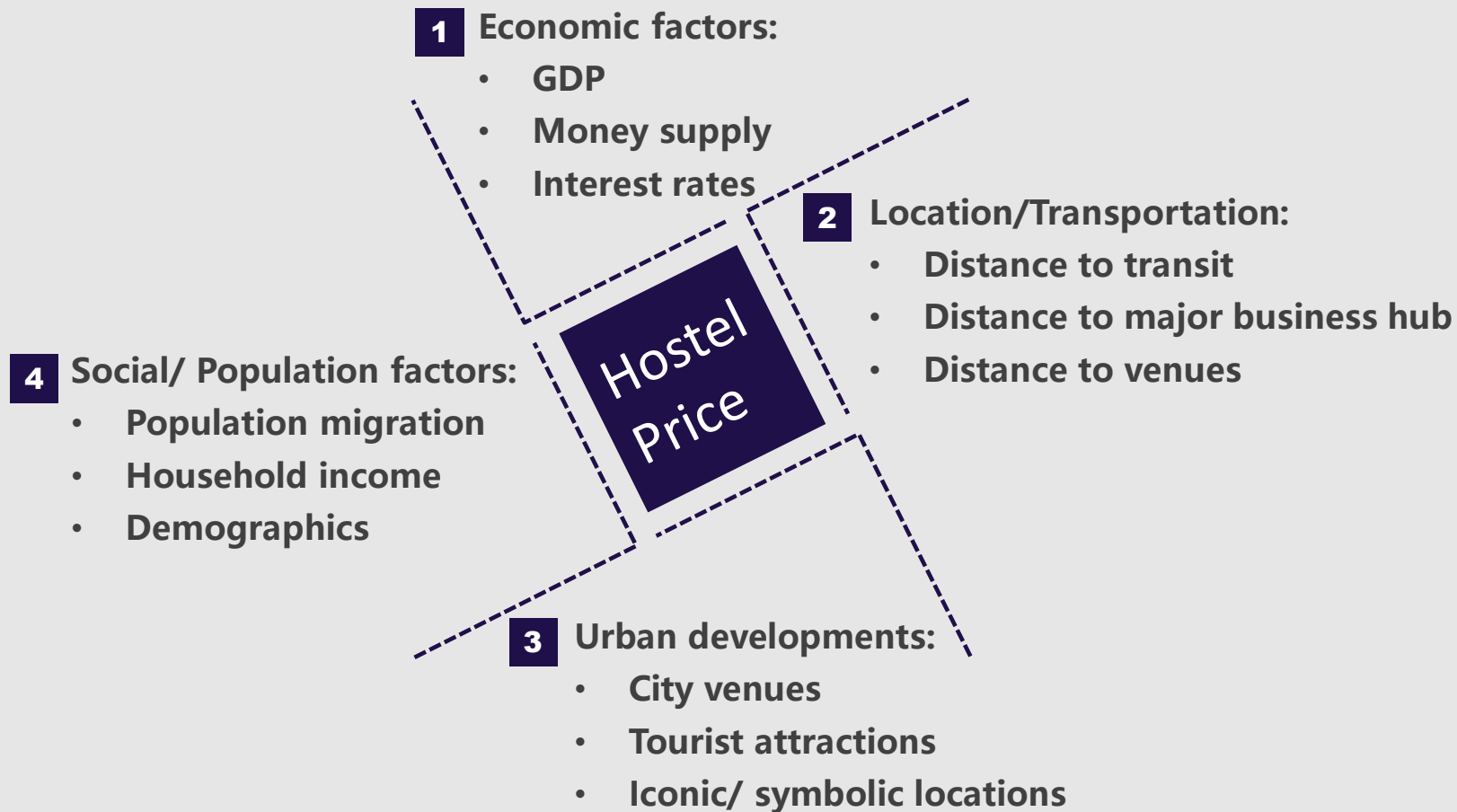
# Tokyo hostel: CPI Analysis

Tokyo hostel industry

The background image shows a dense urban skyline at dusk. A large, curved skyscraper with a grid-like facade is prominent on the left. The sun is low on the horizon, creating a warm, golden glow that reflects off the glass surfaces of the buildings. A semi-transparent dark blue rectangular box is centered over the image, containing the title and introductory text in white.

# Introduction

As of 2020 Tokyo has the highest city GDP in the world, leading New York by a slight margin. Tokyo is also one of the most densely populated city in the world with 37 million people living in an area of 2,193 km<sup>2</sup>. Tokyo also attracts one of the most visitors around the world each year. In this study we will examine how prices vary according to different locations around Tokyo





1. **Tokyo Land price:** Obtained comprehensive list of Tokyo land price index by district from local real estate agency platform (Utinokati).
2. **Hostel neighborhood:** Obtained most common venues of district area using Foursquare API.
3. **Hostel dataset:** Obtained from kaggle.

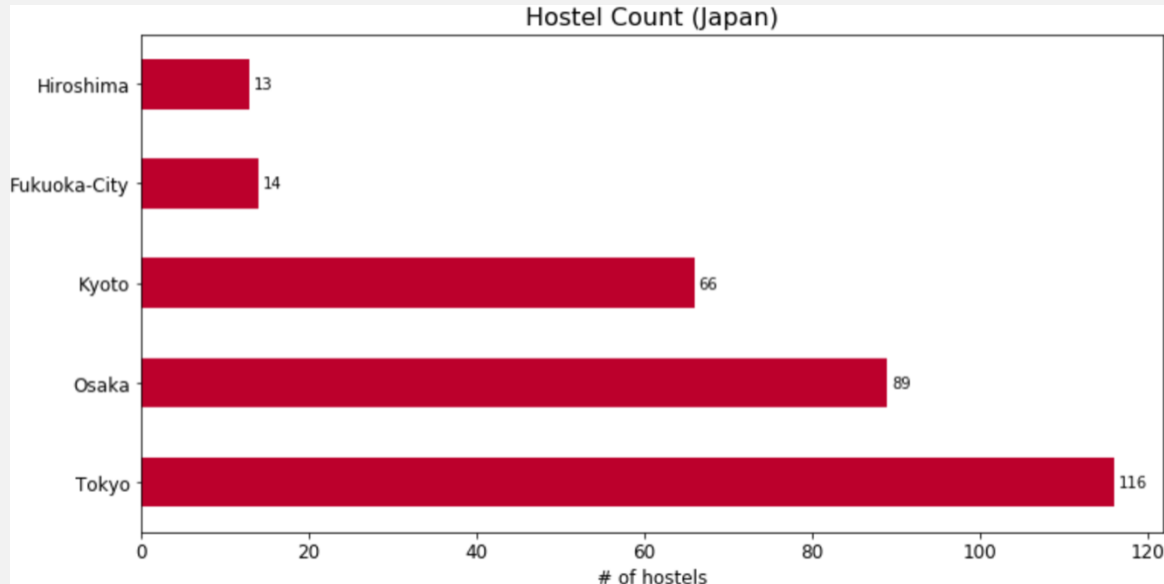
Name	City	StartPrice	DistanceFromCityCentre	OverallScore	RatingCategory	Atmosphere	Cleanliness	Facilities	Location	Security	Staff
"Bike & Bed" CharinCo Hostel	Osaka	3300	2.9	9.2	Superb	8.9	9.4	9.3	8.9	9.0	9.4
&And Hostel Akihabara	Tokyo	3600	7.8	8.7	Fabulous	8.0	7.0	9.0	8.0	10.0	10.0
&And Hostel Ueno	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



1. **Exploratory data analysis:** used to provide useful insights to the reader, travelers and investors. Used the list of hostels from Hostel dataset and use Foursquare API to get venues around the Hostel. I then used EDA to explore the neighborhood and examine its effect on hostel prices
2. **Visualize geography:** Used python folium library to create choropleth map of geographic location and hostel price.
3. **Prescriptive analytics:** Used clustering(K-Means) combined with land price to develop clustering models to predict where a new hostel should be opened.



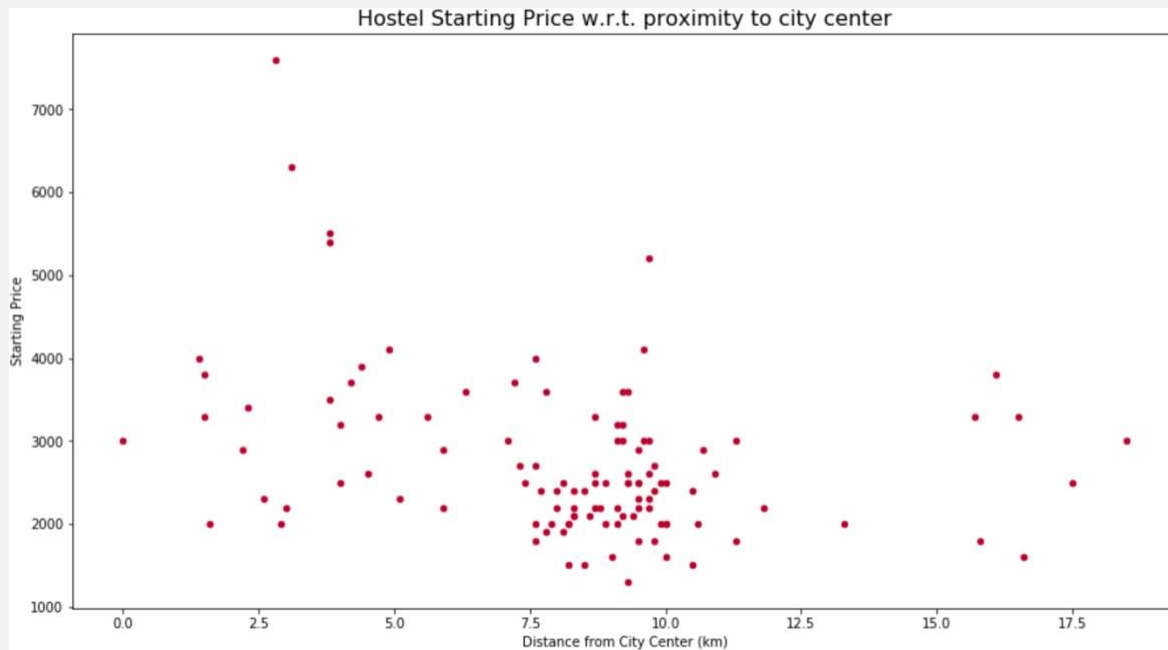
Tokyo has the greatest number of hostels since it is one of the largest cities in the world and attract a lot of visitors. Also, the two airports in Tokyo make it an important hub in the pacific rim.



Next, we visualized all the hostels on map using Folium and Open Street Maps. Below is the geo- visualization of the hostels in Tokyo. As we can see, many hostels are located near the centre with density reducing as we move away from it. In the report we have visualized the same map based on multiple criteria and clusters

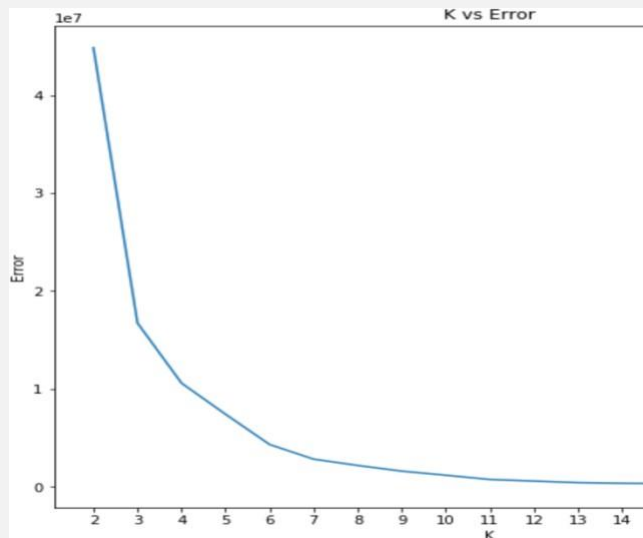


Next, we checked if the distance of a hostel from the city center has any effect on the price. We can see a very weak negative correlation between the price and distance of hostel from city center. Our intuition is supported mathematically by the Pearson coefficient which turned out to be -0.3. The correlation coefficient is very small to make any strong inferences.





We performed clustering twice based on different set of parameters. First, we clustered using the different rating scores, distance from city center, and starting price. We used K-Means clustering algorithm and found out the K by using the elbow method. The K on our case is 6, since the error doesn't decrease much after this point.



After analyzing the data we have a summary of findings listed below which may be useful for travelers or investors:

- The highest number of hostels are located in Taito-ku and Chuo-ku.
- Sumida-ku is an outlier since it is close to Taito-ku and 43% cheaper than Taito-ku.
- The price of hostels does depend 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 transportation or historic site positively affects the hostel rating.
- Price and distance from city center influence hostels rating for value greatly.

In conclusion,

Using existing dataset combined with data collected from Foursquare API and data scraped from the web we have identified potential areas for high CPI hostels. We performed EDA and clustering methodology to arrive at the given solutions. If travelers were to visit Tokyo it would be wise to chose hostels with a high CPI such as those away from city center and those with high security ratings. If travelers are more transit cautious then a hostels in the Chuo-ku area are still viable after weighing in time as a value of money.



# THANKS

Hope to see your visit to Tokyo