

# SINGAPORE

## OPPORTUNITIES FOR INFRASTRUCTURE

---



---

# Introduction

## Background

Singapore is a prime destination for tourists, during both the summers and winters. Singapore attracts over **three times its entire population** in tourists every year, and presents a great opportunity for development of hotels. Currently there are around 500 hotels, with most of them clustered around one area. Thus, there is a great opportunity to build hotels.

## Problem

In order to find the most profitable and ideal locations for building hotels, data about the location and popularity of **tourist attractions and hotels** is required. It is also necessary to figure out the quality of location by analysing the level of development of a locality.

## Data

### Data Source

1. Districts of Singapore, and their geolocation : Districts were web-scraped from **Wikipedia** and corresponding geolocations were obtained from **geocoder** module in python
2. Hotel Information : Obtained from <https://data.gov.sg/dataset/hotels>
3. Tourist Attractions : <https://www.thrilophilia.com/places-to-visit-in-singapore> (Web-Scraped and then geolocations extracted)\
4. Common venues in each district : Foursquare API

---

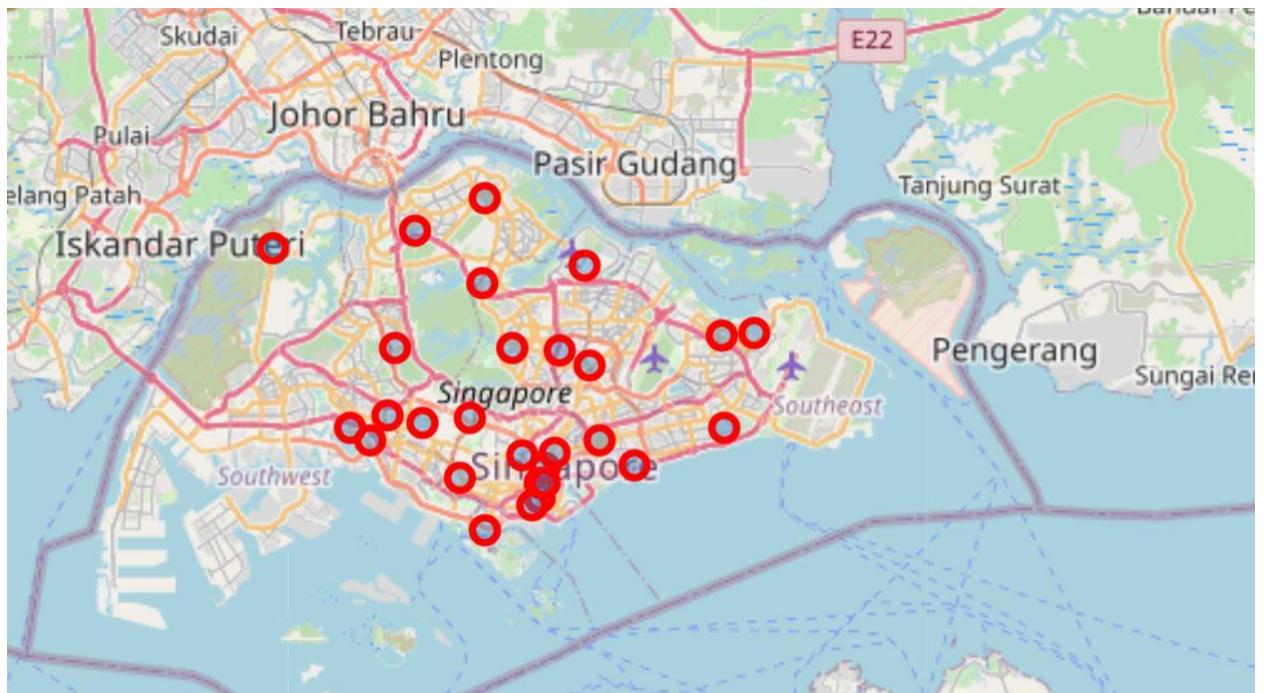
# Methodology

## Data Source

- 1) Analysing different districts to spot less developed areas.
- 2) Finding districts with popular tourist attractions.
- 3) After districts have been identified, finding the missing components to make the district more hospitable.

## Process

- 1) District data was obtained and plotted to get a general idea of the various locations.

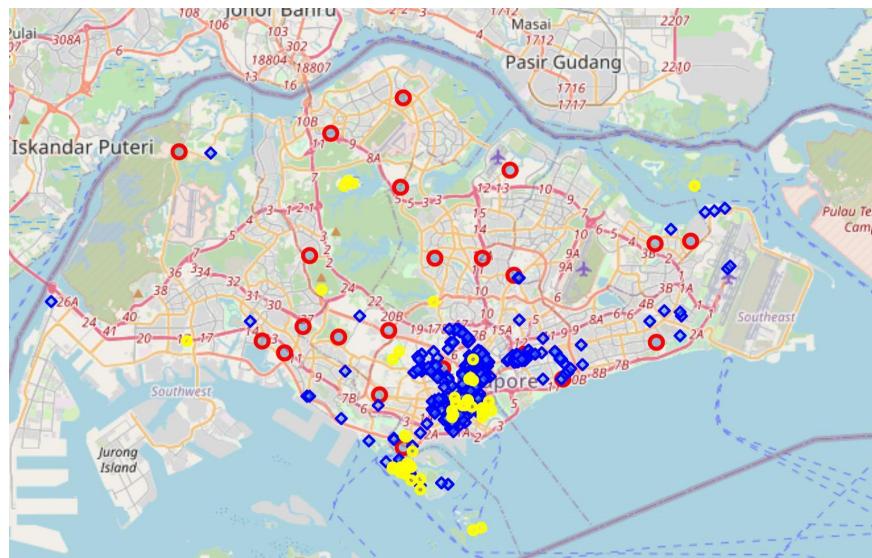


- 
- 2) Hotel Data was added to get an initial idea of which districts are more developed and generally more popular.



Already, we can observe that the southern parts have large clusters of hotels whereas the northern parts have very few hotels.

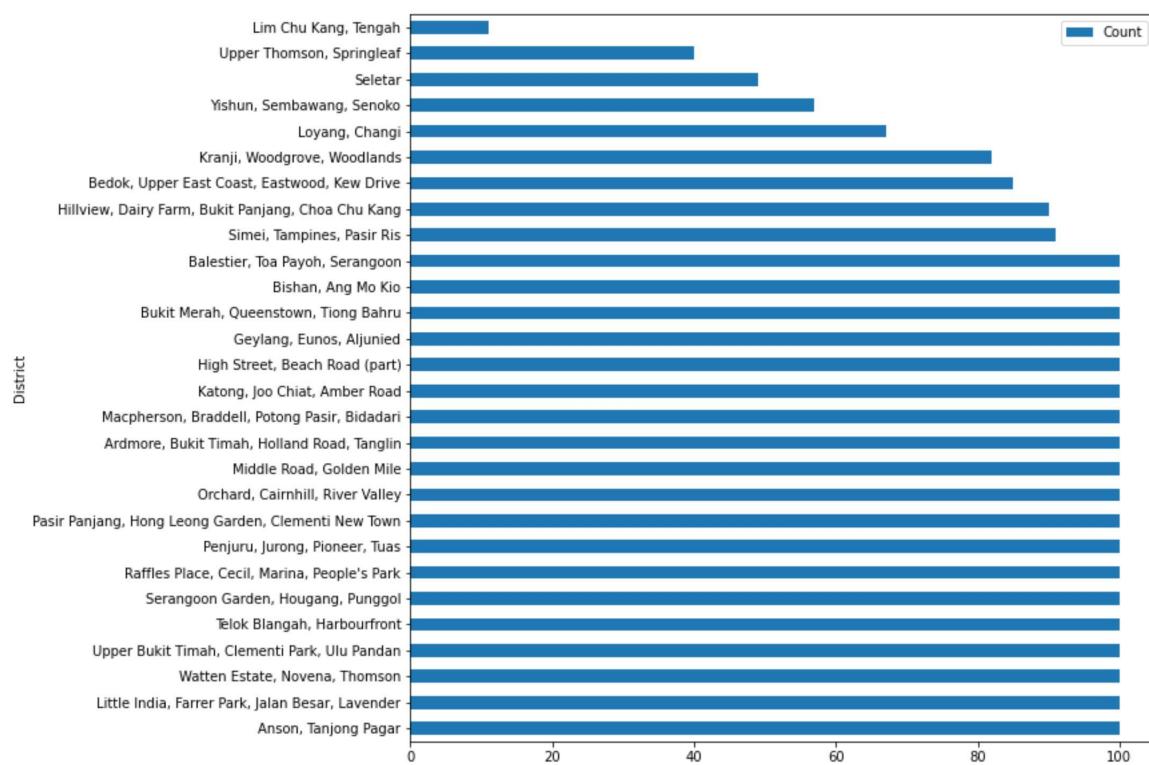
- 3) In order to analyse the reasons for such a distribution of hotels, tourist attraction data was added, which gave a fair idea as to why most hotels are concentrated towards the south of Singapore.



---

We can clearly see that most of the popular tourist attractions are concentrated towards the south, which would be a part of the explanation as to why the hotels were all clustered.

- 4) The next step is analysing each district to see if there is an opportunity. The districts were monitored by the common venues like restaurants and gyms in the area, to understand which areas are less developed.

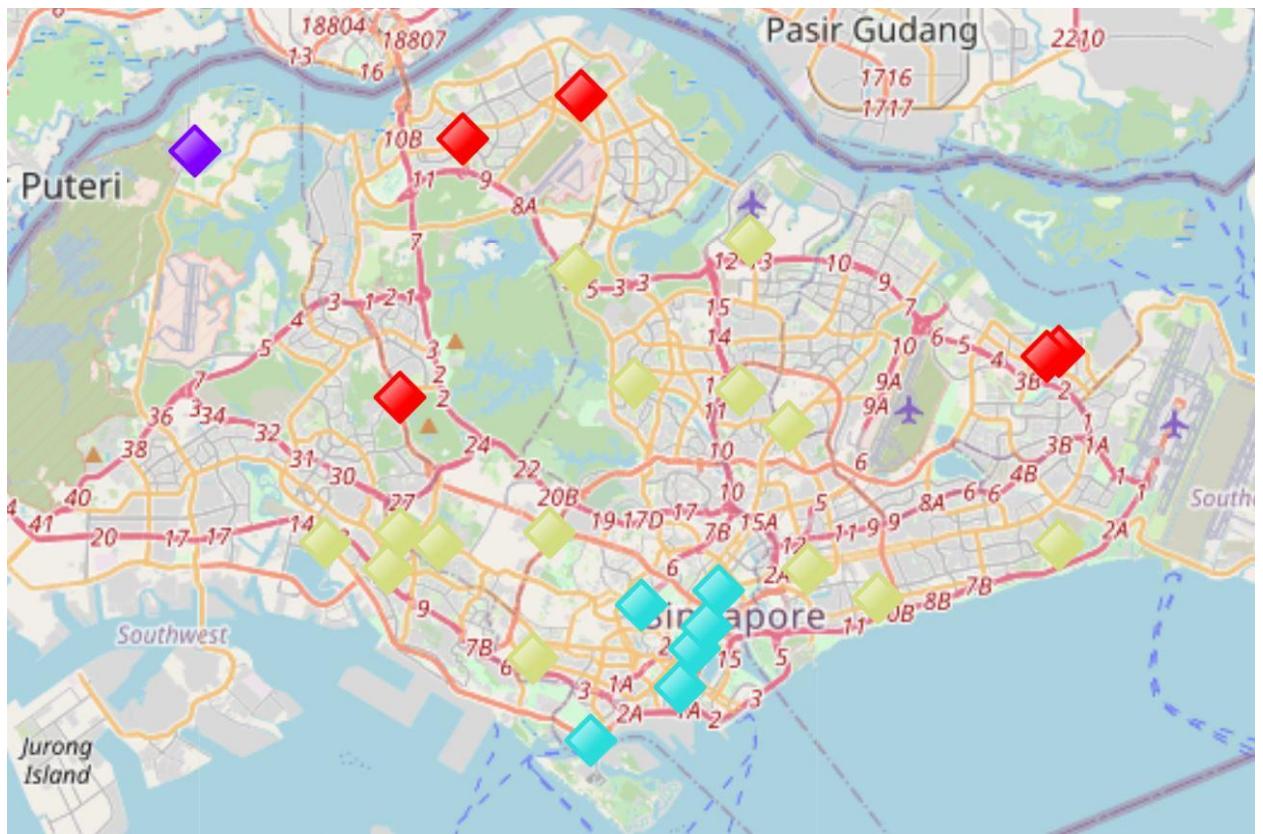


This graph, tallied with the district locations showed that the southern districts have a lot of different venues and there is not much scope for development there.

However, some districts which were not as developed as the others, showed promise. (For example, **Kranji, Woodgrove, Woodlands** and **Hillview, Dairy Farm** districts are enough to have basic amenities as well as certain special facilities but are not developed enough to have many hotels around them.)

- 
- 5) Districts were clustered to observe more similarities and differences between different districts and to observe patterns that might have been missed.

The **K-means Clustering Algorithm** resulted in the following division.



### Green Cluster

THIS IS THE CLUSTER WHICH COVERS A PRETTY LARGE AREA OF SINGAPORE. IT INCLUDES 15 DISTRICTS. WE CAN SEE THAT THE MOST COMMON VENUES IN THESE DISTRICTS INCLUDE RESTAURANTS AND CAFES OF DIFFERENT KINDS, COVERING MANY CUISINES. THUS, IT IS NOT VERY ECONOMICAL TO OPEN ANY RESTAURANTS IN THESE DISTRICTS

---

### Blue Cluster

THIS IS THE CLUSTER LOCATED RIGHT AMONGST THE TOURIST ATTRACTIONS. THIS IS THE AREA WHICH CONTAINS MOST BIG HOTEL CHAINS AND APART FROM RESTAURANTS, ALSO HAS OTHER ATTRACTIONS LIKE ART GALLERIES. THIS AREA IS ALREADY VERY DEVELOPED AND CLEARLY, HAS THE HIGHEST FOOT TRAFFIC IN SINGAPORE.

### Purple Cluster

THIS CLUSTER, ONLY INCLUDES ONE DISTRICT, LOCATED FAR NORTH WEST. IT IS AWAY FROM MOST TOURIST ATTRACTIONS AND IS MAINLY SUBURBAN. IT IS NOT VERY SUITABLE FOR NEW PROJECTS, AS THERE IS VERY LITTLE DEVELOPMENT

### Red Cluster

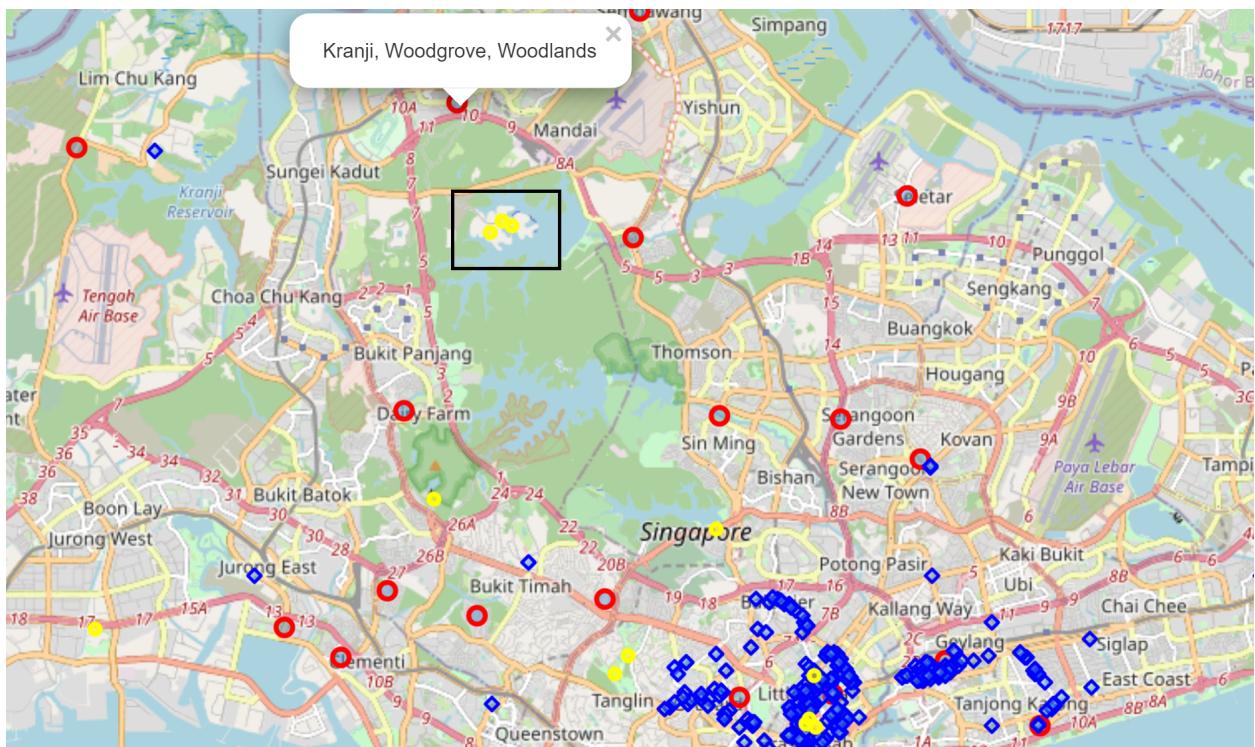
THIS CLUSTER CONSISTS OF DISTRICTS THAT ARE NOT FULLY DEVELOPED (AS SEEN FROM THE EARLIER GRAPH), BUT DEVELOPED ENOUGH TO HAVE SOME GREAT TOURIST ATTRACTIONS LIKE **THE NIGHT SAFARI, SINGAPORE ZOO** AROUND IT.

## Results

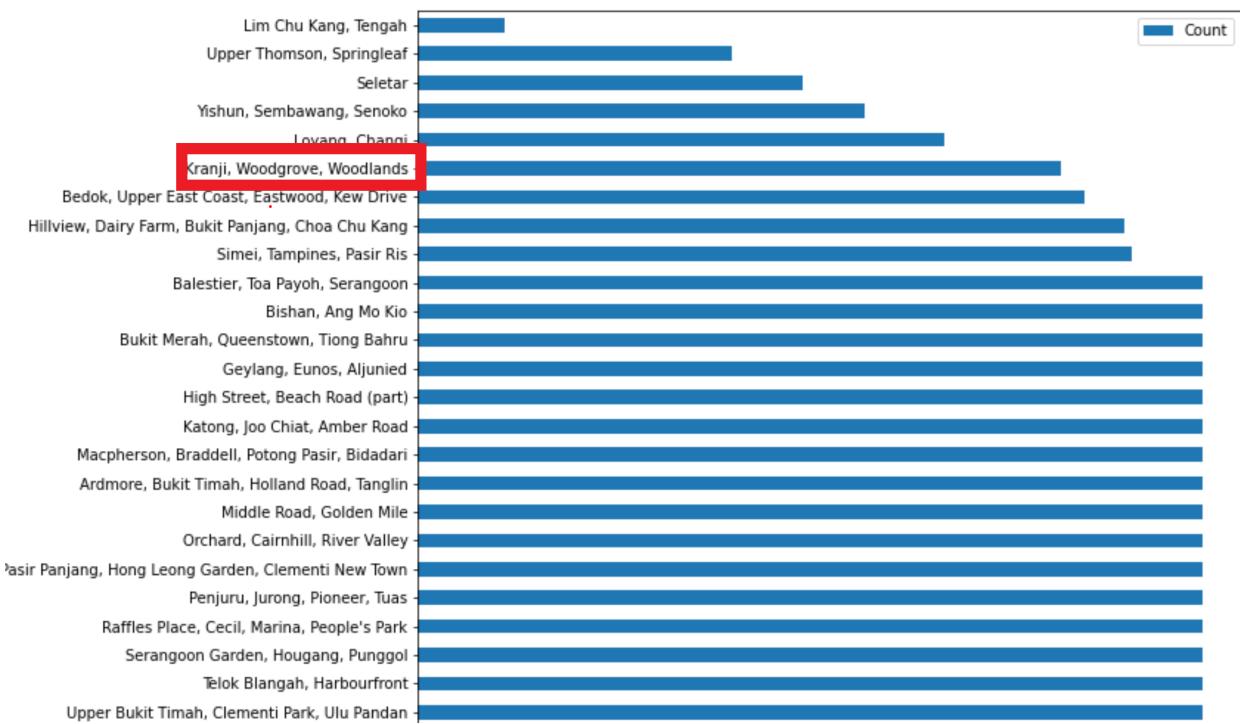
The next step is analysing districts in the red district to find out the one which has a more advantageous location than the others.

District 24, **Kranji, Woodgrove, Woodlands**, was finally picked for the following reasons :

1. It is located in close proximity to 3 very big tourist attractions in Singapore, namely **Night Safari, Singapore Zoo and The River Safari**.



2. It is among the lesser developed districts.



3. Clearly, there is a lack of hotels in this district, and the closest one is in a different district altogether.



---

## Discussion

While gathering and analysing data, I observed why Singapore is such a popular tourist destination.

- 1) The weather is great throughout the year. (Equatorial Climate)
- 2) The entire island is well connected through Metro, Buses and Roads.
- 3) There are attractions for all kinds of people - Art Galleries, Amusement parks, Wildlife Safaris and Zoos.
- 4) There are all kinds of restaurants - Indian, Thai, Sushi, Chinese, Continental, Bars, Cafes etc.

Most of the entertaining activities are located towards the south, due to the close proximity to water. These areas have already been exploited and heavily developed. The rates are also higher and there are more shopping complexes, to maximize profits from tourists.

## Conclusion

This has been a very interesting project, where I have learned a lot about Singapore as a country, and about what attracts tourists.

I feel I have successfully identified a very promising location - **District 24 : Kranji, Woodgrove, Woodlands** for infrastructure opportunity, particularly building hotels.

We started with identifying the districts, then understanding the distribution of hotels and tourist attractions. We then segregated the districts into clusters to identify less developed areas and infrastructure opportunities.