IBM Applied Data Science Capstone

TOPIC: Finding ideal hotels in Singapore

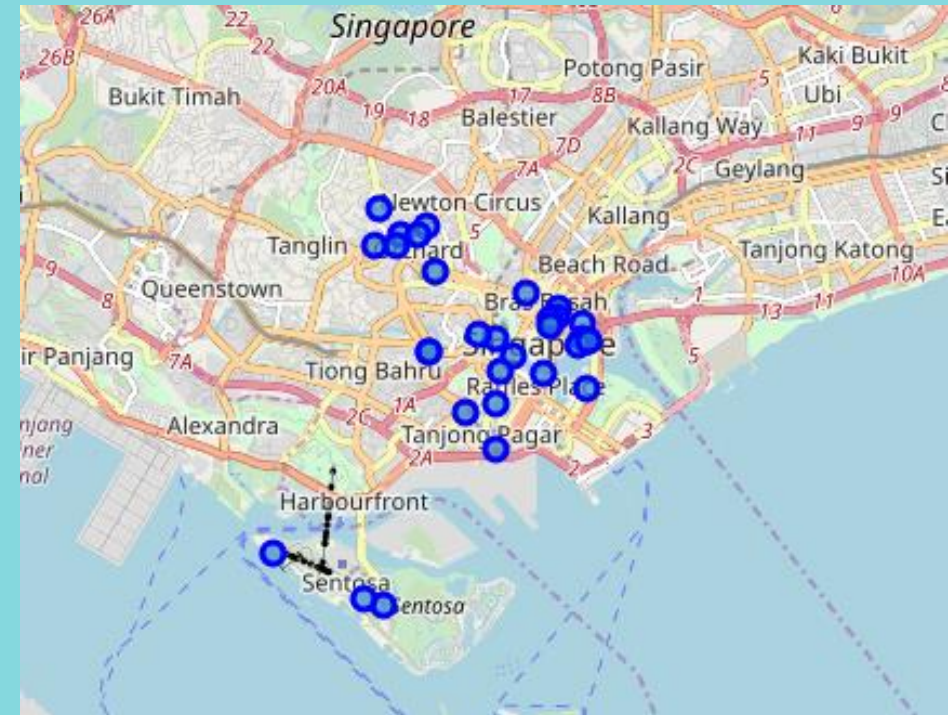By: Do Thi Kim Uyen

Date: 01-12-2020

1. Business problem

➢ As travelers, they should find a hotel which is easy to reach to shopping mall also entertainment area, sight-seeings..

➢ If tourist's pocket has its limitation, they do not want to put much of their expenses on just accommodation.

➢ → In Singapore, if travelers are looking for an ideal hotel to stay, which hotel would you recommend them?

2. Data

➢ **List of hotels in Singapore**

➢ **Latitude and longitude coordinates of those neighborhoods. This is required to plot the map and also to get the venue data.**

➢ **Venue data, particularly data related to hotels. We will use this data to perform clustering on the neighborhoods**

Visualize hotels datasets in Singapore

3. Methodology

01 Define the business problem

02
- Download the data and preprocess using web scarping method
- Explore data and get coordinates by Geocoder package

03 Using Foursquared API to get the top 100 venues surrounds hotels we had

04 Clustering the data by using K-Me ans algorithm

4. Analysis
Analyse Venues

➢ Obtained the number of existing facilities and their type and location in every hotel area with Foursquared API with a limit 100 venues and within 500 meters
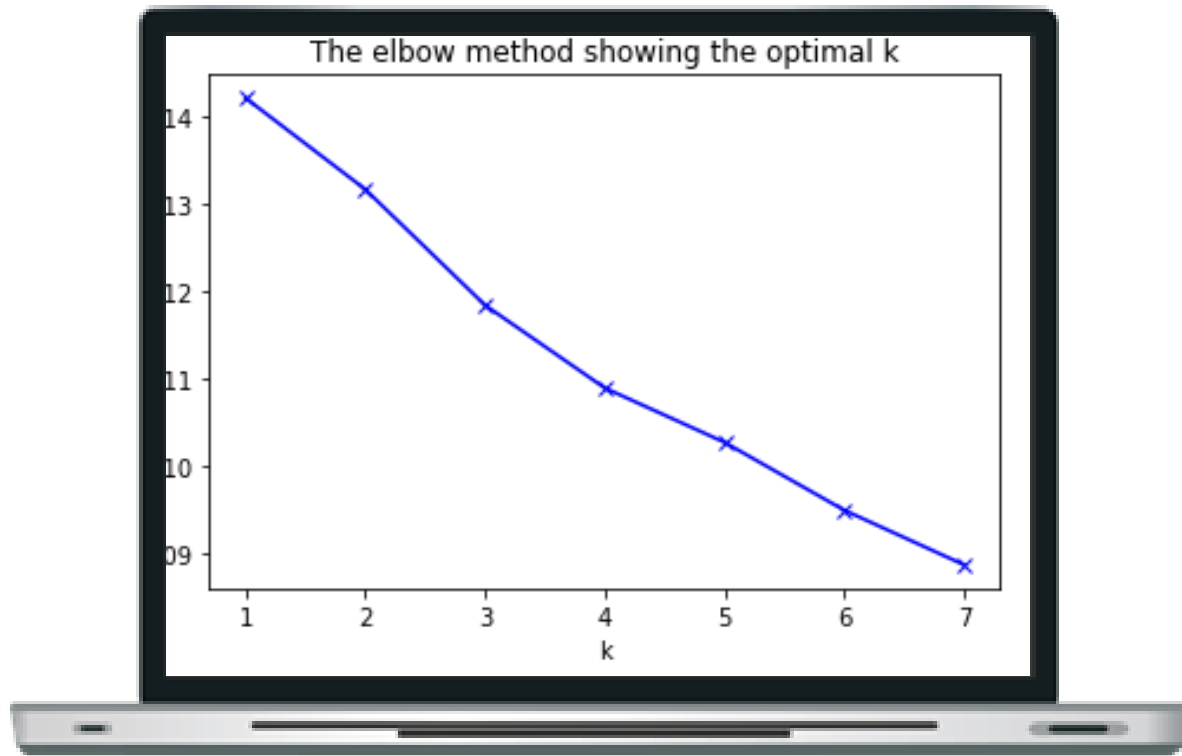
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Naumi Hotel	1.29577	103.85509	JW Marriott Hotel Singapore South Beach	1.294469	103.855793	Hotel
1	Naumi Hotel	1.29577	103.85509	Tom's Palette	1.296079	103.856757	Ice Cream Shop
2	Naumi Hotel	1.29577	103.85509	Raffles Hotel	1.294722	103.854090	Hotel
3	Naumi Hotel	1.29577	103.85509	South Beach	1.294976	103.856496	Shopping Mall
4	Naumi Hotel	1.29577	103.85509	Fairmont Singapore	1.294261	103.853931	Hotel

Analyse Venues

➢ Print each Hotel with the top 10 most common venues and put in a new DataFrame

	Neighborhood	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	Capella Singapore	Theme Park Ride / Attraction	Theme Park	Scenic Lookout	Spa	Resort	Indian Restaurant	Pool	Food Court	Café	Spanish Restaurant
1	Conrad Centennial Singapore	Hotel	Japanese Restaurant	Hotel Bar	Buffet	Café	Bakery	Chinese Restaurant	Coffee Shop	Spa	Lounge
2	Damenlou Hotel	Coffee Shop	Japanese Restaurant	Bakery	Café	Hotel	Ramen Restaurant	Sushi Restaurant	Soup Place	Indian Restaurant	Shopping Mall
3	Fairmont Singapore	Café	Hotel	Japanese Restaurant	Lounge	Shopping Mall	Cocktail Bar	French Restaurant	Chinese Restaurant	Dessert Shop	Concert Hall
4	Garcha Hotels	Japanese Restaurant	Hotel	Nightclub	Bar	Yoga Studio	History Museum	Coffee Shop	Food Court	French Restaurant	Gym

Cluster Neighborhood

The elbow method showing the optimal k

✓ Using Elbow method to define the best k value. From the graph, choosing 5 cluster for optimum k of K-Means.

Cluster **Neighborhood**

✓ Using Folium again to visualise our 5 clusters, which should be a starting point for tourists to explore and search for ideal hotel area for their trip to Singapore.

5. Results and Discussion

Examine Cluster and Discuss

➢ **Cluster 1**. Contains 18 hotels, top 10 common venues mainly conclude Shopping Mall, Restaurant, Barkery, Bar, Buffet,.. This cluster may in the downtown center of Singapore, which has many suitable destination for tourists to explore food, culture and enjoy shopping

	Latitude	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	1.295770	Japanese Restaurant	Café	Chinese Restaurant	Bakery	Clothing Store	Dessert Shop	Shopping Mall	French Restaurant	Lounge
6	1.294650	Café	Japanese Restaurant	Shopping Mall	French Restaurant	Dessert Shop	Coffee Shop	Lounge	Chinese Restaurant	Event Space
7	1.294240	Hotel	Japanese Restaurant	Lounge	Shopping Mall	Cocktail Bar	French Restaurant	Chinese Restaurant	Dessert Shop	Concert Hall
8	1.298490	Hotel	Japanese Restaurant	Restaurant	Art Gallery	Yoga Studio	Bakery	Karaoke Bar	Ice Cream Shop	Bookstore
9	1.293200	Hotel	Shopping Mall	Japanese Restaurant	French Restaurant	Event Space	Cocktail Bar	Lounge	Coffee Shop	Italian Restaurant
10	1.293590	Japanese Restaurant	Hotel Bar	Buffet	Café	Bakery	Chinese Restaurant	Coffee Shop	Spa	Lounge
12	1.283250	Boutique	Theater	Nightclub	Bridge	Noodle House	Japanese Restaurant	Garden	Bar	Tea Room

5. Results and Discussion

Examine Cluster and Discuss

➤ **Cluster 2**. Contains just 1 hotel location, which is near restaurant, resort and golf course… This cluster is suitable for visitor who love to relax and enjoy a soft outdoor activity.

	Latitude	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
25	1.24901	Resort	Spanish Restaurant	Golf Course	Chinese Restaurant	Restaurant	Hotel	Italian Restaurant	Hotel Bar	Library

5. Results and Discussion

Examine Cluster and Discuss

➤ **Cluster 3**. Also contains only 1 hotel locations and mainly surround by restaurants. Any travelers enjoying in food from many cultures can choose to be here for their trip, as they are easy to reach any restaurants nearby their hotels

	Latitude	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
22	1.311561	Indian Restaurant	Steakhouse	Italian Restaurant	Lounge	Dim Sum Restaurant	Dance Studio	Chinese Restaurant	Cantonese Restaurant	Seafood Restaurant

5. Results and Discussion

Examine Cluster and Discuss

➢ **Cluster 4**. Contains only 2 hotels, surround by park, drink stores, resort, spa.. same as cluster 2. This cluster may be a good choice for tourist who love relaxing and soft and slow activities.

	Latitude	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
24	1.257012	Bar	Historic Site	Theme Park	Resort	Restaurant	Breakfast Spot	Café	Coffee Shop	Hot Dog Joint
26	1.249820	Theme Park	Scenic Lookout	Spa	Resort	Indian Restaurant	Pool	Food Court	Café	Spanish Restaurant

5. Results and Discussion

Examine Cluster and Discuss

➢ **Cluster 5**. Contains 8 hotels locations with many utilities nearby like Restaurant, coffee shop, museum, yoga and gym, also nightclub and spa,.. same as cluster 1, this cluster may be around downtown Singapore, which related many entertainment places for active tourists.

	Latitude	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	1.273698	Japanese Restaurant	Bakery	Café	Hotel	Ramen Restaurant	Sushi Restaurant	Soup Place	Indian Restaurant	Shopping Mall
2	1.279323	Chinese Restaurant	Coffee Shop	Café	Japanese Restaurant	Korean Restaurant	Food Court	Dessert Shop	Cocktail Bar	Asian Restaurant
3	1.280780	Coffee Shop	Cocktail Bar	Chinese Restaurant	Food Court	Bar	Restaurant	Beer Garden	Japanese Restaurant	Dessert Shop
4	1.288350	Hotel	Nightclub	Bar	Yoga Studio	History Museum	Coffee Shop	Food Court	French Restaurant	Gym
5	1.285880	Hostel	Japanese Restaurant	Seafood Restaurant	Chinese Restaurant	Café	Yoga Studio	Vegetarian / Vegan Restaurant	Spa	Miscellaneous Shop
11	1.285610	Chinese Restaurant	Gym / Fitness Center	Waterfront	Italian Restaurant	Japanese Restaurant	Salad Place	Bar	Café	Gym
28	1.291260	Hotel	Bar	Spa	Café	Ramen Restaurant	Nightclub	Seafood Restaurant	Bakery	Supermarket
29	1.291810	Hotel	Nightclub	Bar	Bakery	Café	Spa	Pub	Coffee Shop	Ramen Restaurant

6. Conclusion

01

➤ The Foursquared location data was leveraged to compared each hotel location to provide believable suggestions for tourists. With the help of K-Means Clustering algorithm., all hotels were clustered into 5 groups. And each group's advantages can be clarified easily

02

➤ Further analysis can be done with these 5 cluster, but may use additional dataset related to Singapore facilities to explore more in details based on tourists' needs

Thank you
for reading!