

Coursera Capstone Project

The Battle of Neighborhoods

By Cheung Ngai Lam Irene

December 23, 2019Result

Table of Contents

Introduction.....3

Business Problem.....3

Data.....3

Methodology4

Result.....6

Discussion6

Conclusion6

Introduction

Singapore, there are a lot of tourist visit all year around. They like to do a lot of sightseeing around the city. After a long day, tourist would like to take a rest and have a cup of coffee to relax in this beautiful place. Therefore, it is an opportunity to open coffee shop.

If someone is looking to open a coffee shop in Singapore, the question is where would you recommend them to open it? The business problem is that in order for a coffee shop to be profitable, there must be enough customers. It is not worth to open a new coffee shop in the immediate proximity of existing ones.

Business Problem

The objective of this project is to analyze and select the best locations in Singapore to open a new coffee shop. Using data science methodology to provide business solution.

Let's also make sure that audience is explicitly defined to be the local restaurant entrepreneur in Singapore and they care about this problem because the location of the new coffee shop has the significant impact on its revenue and they want this to be a successful one.

Data

The data used to solve this problem is geolocation data collected from FourSquire.com. Foursquare has one of the largest database of 105+ million places and used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the coffee shop category in order to help us to solve the business problem. Adequate explanation and discussion of the data is the following. Data is a single dataframe, containing at least a location of the coffee shop. Explanation of the location data is a standard tuple (lat, lng), where lat stands for latitude and lng for longitude. Some other metadata such as name, postal code, and etc., are also collected. Example of the data used in analysis is shown in Table 1.

No.	Name	Short Name	Address	Postal Code	Latitude	Longitude
1	PPP Coffee	Coffee Shop	#02-19 Funan (107 North Bridge Road)	179095	1.291628	103.849741
2	Starbucks Reserve Store	Coffee Shop	#02-01/02/03 The Fullerton Waterboat House	049215	1.287049	103.853600
3	RONIN	Café	17 Hongkong St	059660	1.287708	103.847177
4	Hoshino Coffee	Café	#B2-55 Capitol Piazza (13 Stamford Road)	178905	1.292988	103.851441
5	The Stamford Brasserie	Café	2 Stamford Rd	178882	1.292836	103.853693

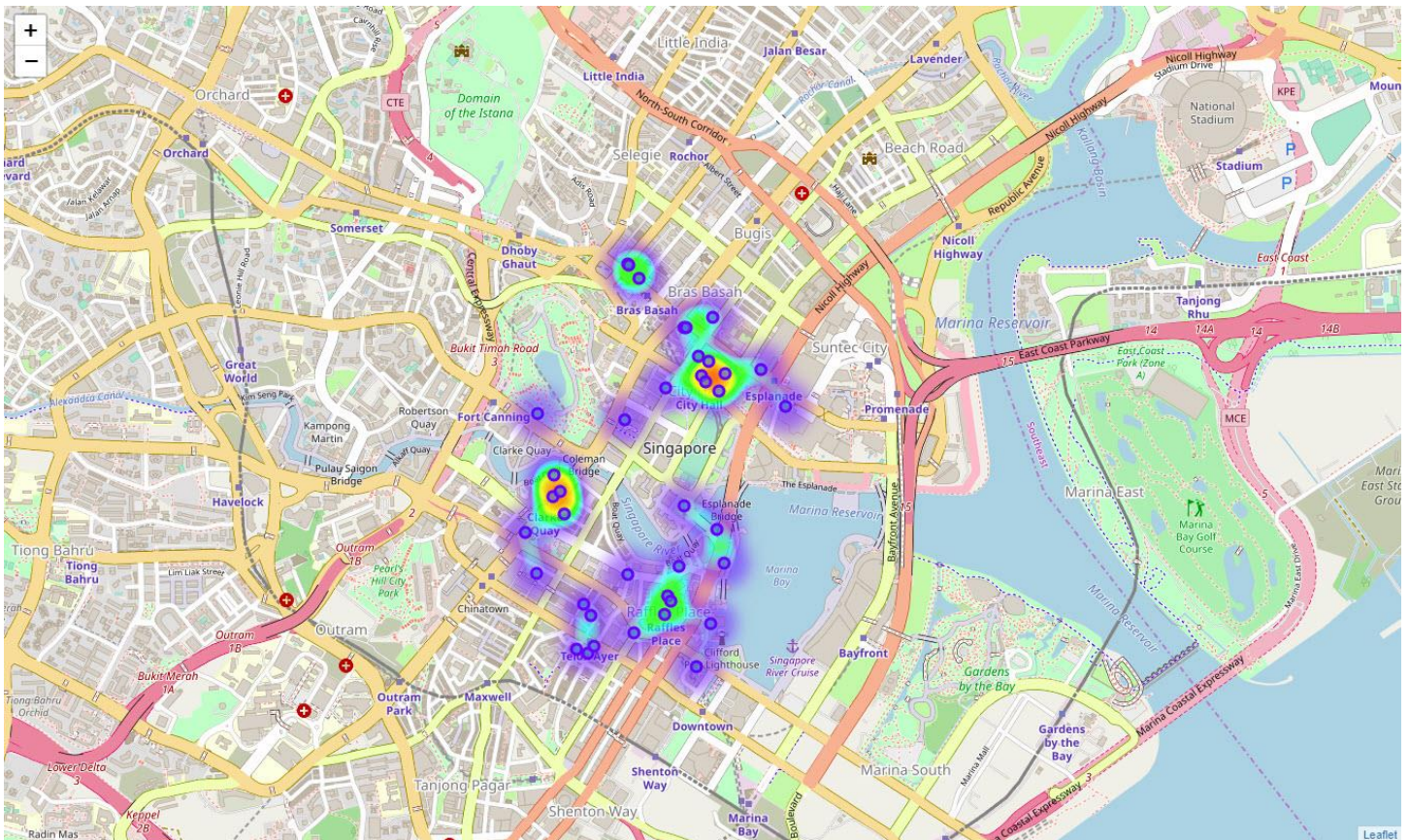
Table 1: First five rows of data used in the analysis.

Data will be used as follow – by knowing the locations of the already existing coffee shops. It is possible to apply unsupervised learning technique like kernel density estimation (KDE) to determine the area of influence of the existing coffee shops.

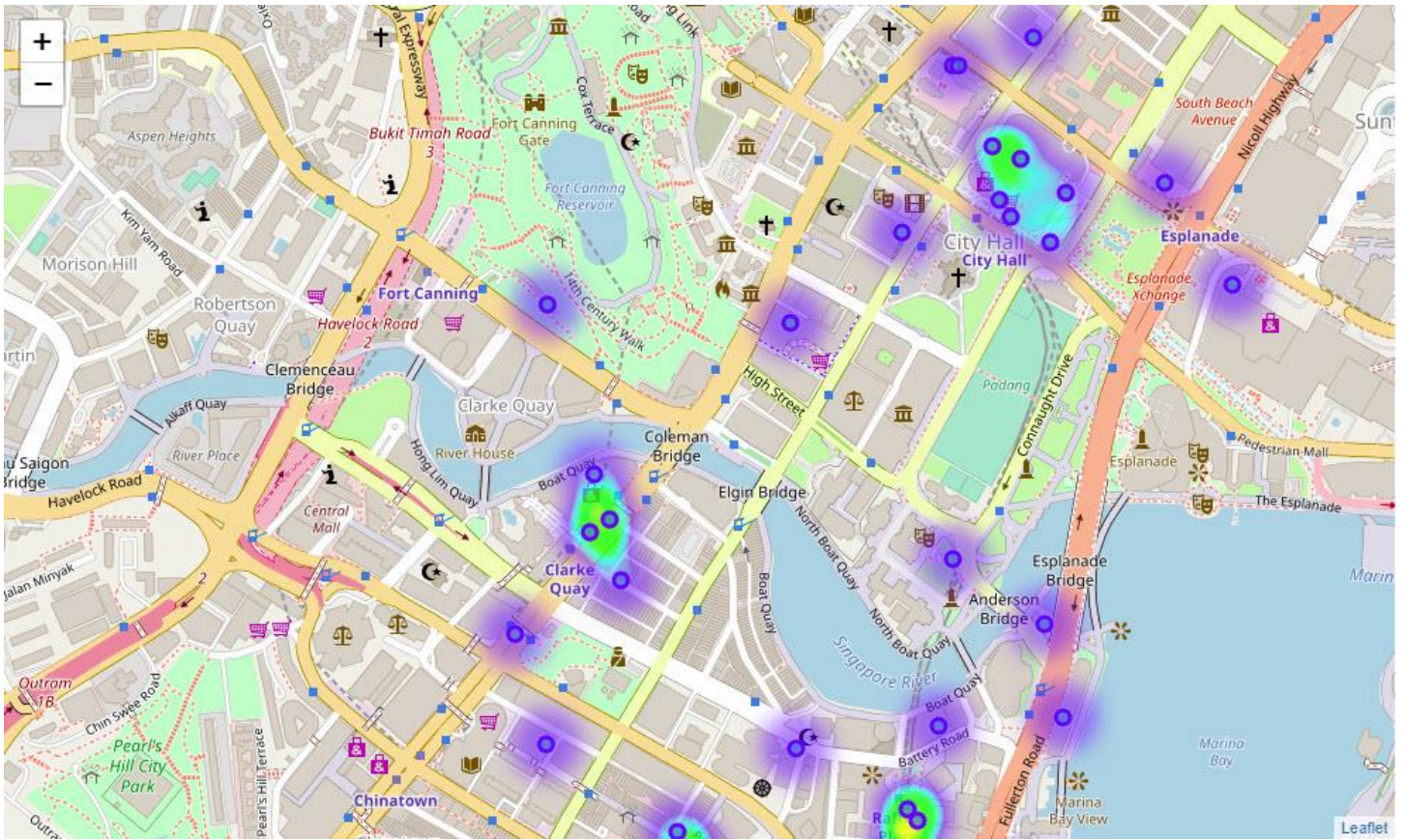
Methodology

We will use Foursquare API to get the top 50 venues that are within a radius of 1000 meters. We need to have a Foursquare Developer Account in order to obtain the Foursquare ID and Secret Key. We then make the API calls to Foursquare passing in the geographical coordinates of the neighborhoods. Foursquare will return the venue data in JSON format and we will extract the venue name, venue category, venue address, venue latitude and venue longitude. With the data, we can filter the “Coffee Shop” as venue category for the neighborhoods.

Also, Heatmap-based kernel density estimation was being used. Heatmap was already implemented as plugin for Folium which was used to visualize data on the map. Visualization is shown in map 1 and map 2.



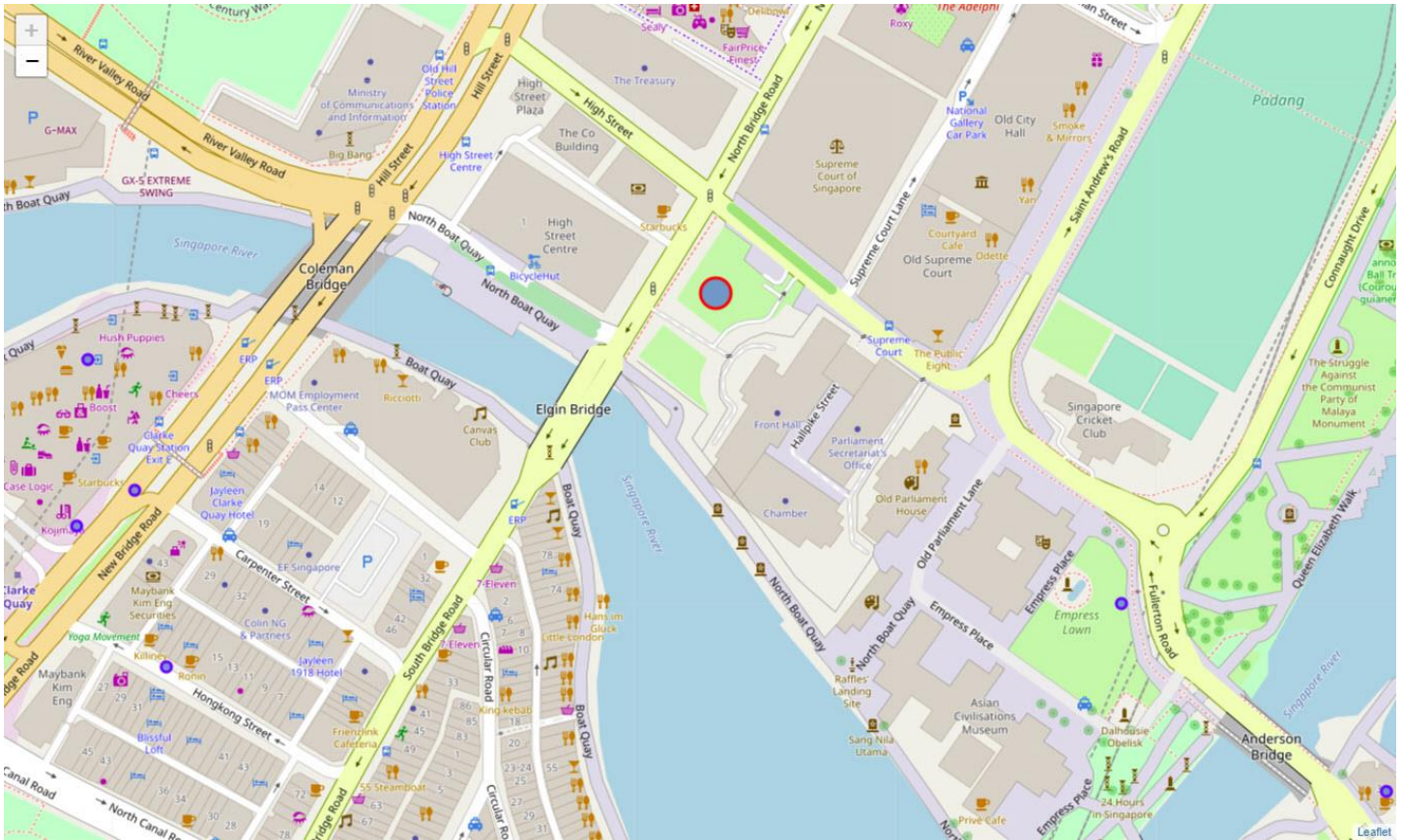
Map 1: Data visualized to the map of Singapore, including heatmap-based kernel density estimation.



Map 2: The zoom in version of the Data visualized to the map of Singapore, including heatmap-based kernel density estimation.

Result

Based on the preliminary result, one possible good location for the new coffee shop would be in crossroad of North Bridge Road and High Street, which is inside the park near the Parliament of Singapore, is shown in map 3.



Map 3: Recommended location for the new coffee shop.

Discussion

Compare with the map in the Result section and the heatmap-based kernel density estimation in the Methodology section, you see that most of the coffee shops are concentrated near the City Hall of Singapore. There is nearly no coffee shop near the Singapore River. This represents a great opportunity and high potential area to open a new coffee shop as there is very little to no competition from existing coffee shop. And tourist can enjoy a cup of coffee and relax a bit after a long day of sightseeing in this beautiful place.

Conclusion

The optimal location for the new coffee shop would be in crossroad of North Bridge Road and High Street, which is inside the park near the Parliament of Singapore. It was estimated based on the data from FourSquare.

It is recommended to the management team to re-run this data science program to get the updated result and use the result into consideration as part of the business growth plan in selecting the new neighborhood to offer their new coffee

shop. This is critical not only to make sure that they got the updated result for better decision making, but also to make sure that they can revalidate the finding from this project.