

Identify potential location for a new Family Mart store

Gift Metavarayuth

Introduction

Family mart is a Japanese convenient store franchise chain. It is planning to expand its franchise to Bangkok where 7-Eleven is the leader in convenient store in the area. In order to find potential locations to open Family Mart stores, Kmeans cluster was used to segment and cluster neighborhoods in Bangkok based on nearby venue categories. Then the neighborhoods which have lower numbers of 7-Eleven store in similar cluster was selected for further explore.

Data acquisition and cleaning

1. Data source

Data required for this project are Bangkok neighborhoods, nearby venues in the neighborhoods, and number of 7-Eleven stores in each neighborhood. The neighborhood data contained neighborhood name and their location was obtained from Wikipedia, while nearby venues and number of 7-Eleven stores in each neighborhood was explored and searched by using Fourquare API.

In order to accurately search for numbers of 7-Eleven store in a neighborhood, the radius variable should be specified in Foursquare request. Therefore, radius was calculated from area of each neighborhood. The area data was obtained from Bangkok database website.

2. Data cleaning

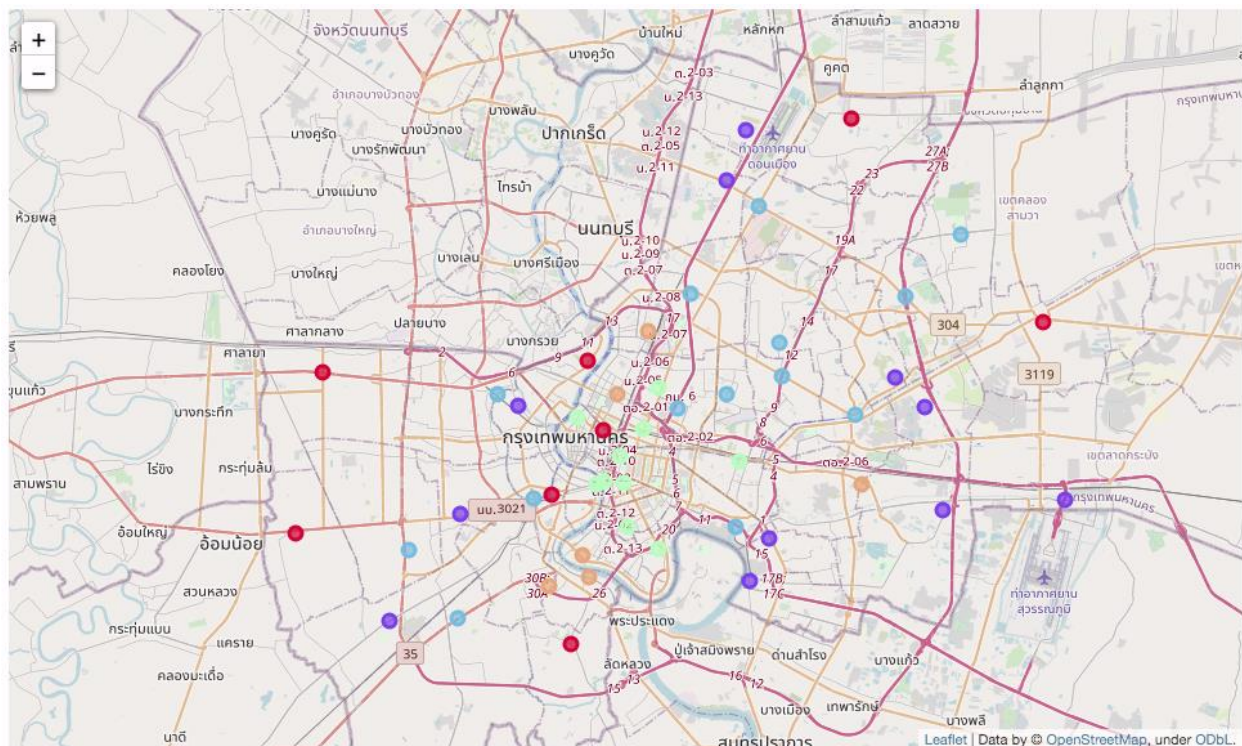
The neighborhood data from Wikipedia and area data from Bangkok database website both have a common column which is neighborhood name in Thai. However, the neighborhood Thai names in area dataframe has prefix “เขต” in them. Therefore, data cleaning was performed to remove the prefix before the two dataframes can be merged together correctly.

After the dataframes were merged, one neighborhood, Rat Burana was missing. The neighborhood was added back to the merged dataframe before the study was proceeded to the next step.

Methodology

After location data and radius values were obtained for each neighborhood, Foursquare API was used to acquire nearby venues and number of 7-Eleven stores in each neighborhood. After that, one hot coding was used to obtained average value of each venue category in the neighborhood. Then, Kmeans cluster analysis was performed to segment neighborhood in Bangkok based on nearby venues. Finally, each cluster was analyzed to identify theme of the cluster and potential location of new a Family Mart store can be recommended based on insight information from the cluster analysis together with number of the competitive convenient store, 7-Eleven in the neighborhood.

Results and discussion



The Kmeans cluster approach segment Bangkok neighborhoods into 5 clusters

1. Center of convenient stores, Thai restaurants and noodle houses (Red dots)
2. Mixed area of coffee shop, convenient store and Thai restaurant (Purple dots)
3. Heaven for noodle lovers and coffee shop hoppers (Blue dots)
4. Tourist district (Green dots)
5. Hotpot hotspot (Pink dots)

The analysis exhibits a very interesting information such as Tourist district that would not be identified if the Kmeans cluster was not done. In addition, although the segmentation was done based on nearby venues in each neighborhood, the neighborhoods in each clusters are similar

to each other in term of distance from city center as well. In general, neighborhood in cluster 4 (Tourist district) are located in the middle of/close to city center. Cluster 3 and 5 are the next area located a little further out of the city center than cluster 4. Cluster 2 and then 1 are the next area that locate further away from city center.

More importantly, out of these 5 clusters, convenient stores seem to be lacking in the tourist district. In addition, Samphanthawong neighborhood in tourist district has the lowest number of 7-Eleven store. Therefore, it is identified as a potential location for opening a new Family Mart store based on this analysis.

Conclusion

In this project, I utilize Kmeans cluster to segment and cluster neighborhoods in Bangkok based on nearby venue categories to identify potential location for a Family Mart store. The Kmeans cluster approach segment Bangkok neighborhoods into 5 clusters which are 1. the center of convenient stores, Thai restaurants and noodle houses, 2. mixed area of coffee shop, convenient store and Thai restaurant, 3. the heaven for noodle lovers and coffee shop hoppers, 4. tourist district, and 5. hotpot hotspot. Out of these 5 clusters, convenient stores seem to be lacking in the tourist district especially Samphanthawong neighborhood. Therefore, it is nominated to be the best location for a Family Mart store based on this study.

Further studies could be conducted by pinpoint a smaller area for the new Family Mart store by analysing sub neighborhood in Samphantawong neighborhood. It is also possible to take into account the number of population per convenient store in each neighborhood.