Feasibility Study of Opening New Korean Restaurant in Jakarta, Indonesia

Bangkit Dana Setiawan

1. INTRODUCTION

Currently Korean plays significant role in global business. I believe that in every country in the world at least ever heard of Korean's entertainment Industries, like K-Pop, K-Drama, and K-Movies. Especially in my country, Indonesia, started with the influence of the Korean entertainment industries, the several communities who idolize this kind of entertainment started to developed. In fact, this influence expands to not only on entertainment industries, but also to other aspects like language, culture and food. This creates new business opportunity for establishing business which focuses on Korean related. For instance, currently the Korean food industry is growing really fast in Indonesia. Almost every city has the Korean Restaurant. Therefore, this is a good timing for business man to involve in this Korean food industry in Indonesia.

Jakarta as the capital city of Indonesia shows the rapid economic growth in the last few years. This means that it will be great chance to open the new business like Korean Restaurant in here. However, opening new restaurant in Jakarta should have several considerations, especially about the competitiveness index. The businessman has to choose the best place in Jakarta to open the new restaurant. The location of the new business will determine whether the business will be a success or a failure.

a. Business Problem

The objective this capstone project is to analyze and select the best location in the Jakarta, as the capital city in Indonesia, to open a new Korean Restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solution to answer business question:

If there is a businessman who is looking the location to open the new Korean Restaurant, what is your recommendation?

b. Target Audience of this project

This project is particularly useful to property developers and investors looking to open or invest in new Korean Restaurant in Jakarta, Indonesia.

2. DATA

These are the following data and also the method to extract them which will be needed to solve the problem:

- List of neighborhoods in Jakarta.
    - This defines the scope of this project which only focus on the city of Jakarta, the capital city of Indonesia.
    - The neighborhoods data in Jakarta will be taken from Wikipedia page https://en.wikipedia.org/wiki/Category:South_Jakarta
    - This contains of 63 neighborhoods
    - This data will be utilized through web scrapping techniques to extract the data from Wikipedia page with the help of Python requests and beautifulsoup packages

- Latitude and Longitude coordinates of those neighborhoods.
    - This will be required to plot the map and also get the venue data
    - After data from wikipedia already extracted, then latitude and longitude of the neighborhoods will be determined using the Python Geocoder package

- Venue data through Foursquare API, especially related with Korean Restaurant.
    - This will be used for performing clustering in the Neighborhoods
    - Foursquare API will provide many categories of the venue data, we are particularly interested in the Korean Restaurant category in order to help us to solve the business problem.

This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium).

3. METHODOLOGY

1. List of data neighborhoods in Jakarta, Indonesia which was withdrawn from Wikipedia page https://en.wikipedia.org/wiki/Category:South_Jakarta then being extracted with technique web scraping using python requests and beautifulsoup packages. This resulted list of names of neighbourhoods data in Jakarta.
2. In order to get geographical coordinates in the form of latitude and longitude from the extracted list of neighborhoods in Jakarta from Step 1, Geocorder package is used. This Geocorder packages allows us to convert address into geographical coordinates with result of latitude and longitude for each neighborhoods in Jakarta.
3. All of the latitude and longitude data for each neighborhoods in Jakarta was gathered and put it into pandas Dataframe.
4. The Daftarame from Step 3 then visualized into map using Folium Packages. The map visualization will be used to perform check and make sure that the geographical coordinate data from Step 2 were correctly plotted in the city of Jakarta, Indonesia.
5. It is needed to register a Foursquare Developer account for obtaining Foursquare ID and Foursquare secret key before use Foursquare API
6. Foursqare API was used to get the 100 venues that are within a radius of 2000 meters. This number was determined to get the representative Venues data to be analyzed next.
7. Foursquare API call was performed to do Foursquare passing in the geographical coordinates of the neighborhoods in a Python loop. This step resulted the venue data in JSON format.
8. The venue data in JSON format from Step 7 then was extracted based on venue name, venue category, venue latitude, and longitude. These data will be used to know how many venues were returned for each neighborhood and determine how many unique category can be curated from all the returned venues.
9. Then it was performed analysis of each neighborhood by grouping the rows by neighborhood and extracted the mean of the occurrence frequency of each venue category. This was also performed to prepare clustering.
10. Since this project goal is to open new Korean Restaurant, then the data will be filtered with only “Korean Restaurant” as venue category for each of neighborhood in Jakarta.

11. Then it was performed clustering on the data by using k-Means clustering with the source of filtered data from step 10. K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroid as small as possible.
12. It was performed clustering the neighborhoods into 3 clusters based on their frequency of occurrence for "Korean Restaurant". The results will allow us to identify which neighborhoods have higher concentration of Korean Restaurant and which have fewer. Based on the occurrence of Korean Restaurant, it will help us to answer the objective of this project as to which neighborhood are most suitable to open new Korean Restaurant.

4. RESULTS

a. Dataframe of Neighborhoods in Jakarta with Latitude and Longitude

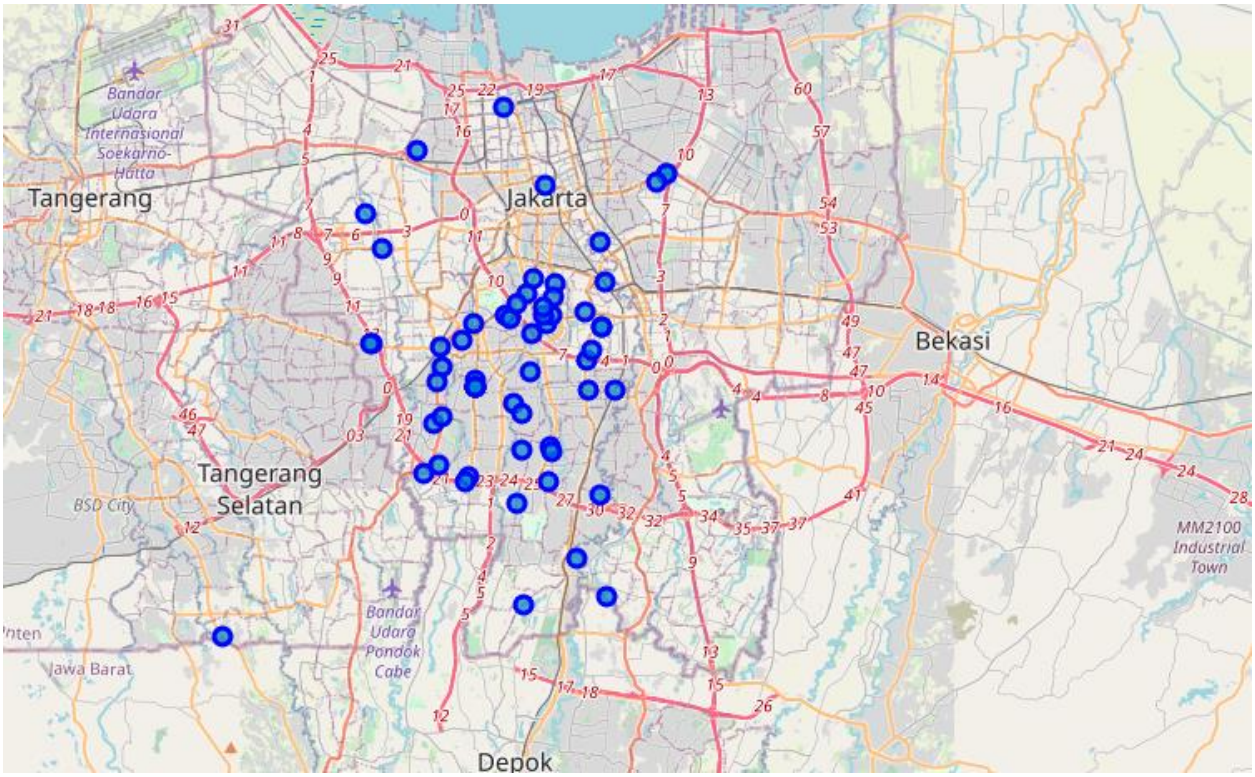
Below is the dataframe of Neighborhoods in Jakarta with Latitude and Longitude after extracted from Wikipedia data using using python requests and beautifulsoup packages and also covertion of Geocoder packages:

]:

	Neighborhood	Latitude	Longitude
0	South Jakarta	-6.25374	106.79731
1	AIS Indonesia	-6.27824	106.82858
2	ASEAN MRT station	-6.16624	106.87657
3	Atma Jaya Catholic University of Indonesia	-6.19417	106.84914
4	Balai Sarbini	-6.25374	106.79731
...	...	...	...
58	Tanjung Barat railway station	-6.29827	106.84905
59	Tebet Honda Park	-6.23891	106.84519
60	Tebet railway station	-6.22939	106.84971
61	Tebet, South Jakarta	-6.22939	106.84971
62	The Plaza Semanggi	-6.21959	106.81440

b. Map Visualization of Neighborhoods in Jakarta

Below is the map visualization of Neighborhoods in Jakarta using the help of folium packages using the dataframe latitude and longitude



c. 100 Venues within radius 2000 meters in Neighborhoods

Below is the data which shows 100 Venues within radius 2000 meters in Neighborhoods which is obtained by using foursquare API

	Neighborhood	Latitude	Longitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	South Jakarta	-6.25374	106.79731	Nasi Uduk Pak Joko Putro	-6.254641	106.796036	Indonesian Restaurant
1	South Jakarta	-6.25374	106.79731	Sopo Ngiro Japanese Food	-6.251255	106.797441	Japanese Restaurant
2	South Jakarta	-6.25374	106.79731	Chung Gi Wa 청기와	-6.252591	106.800207	Korean Restaurant
3	South Jakarta	-6.25374	106.79731	Nasi Campur Kenanga	-6.253358	106.800189	BBQ Joint
4	South Jakarta	-6.25374	106.79731	All Fresh	-6.249712	106.797916	Fruit & Vegetable Store
...	...	...	...	...	...	...	...
5921	The Plaza Semanggi	-6.21959	106.81440	Okuzono	-6.233408	106.812024	Japanese Restaurant
5922	The Plaza Semanggi	-6.21959	106.81440	Fairmont Jakarta	-6.225578	106.799239	Hotel
5923	The Plaza Semanggi	-6.21959	106.81440	Loewy	-6.227236	106.826087	Bistro
5924	The Plaza Semanggi	-6.21959	106.81440	PEPeNERO	-6.215197	106.821084	Italian Restaurant
5925	The Plaza Semanggi	-6.21959	106.81440	Legend of Noodles (취강해물짬뽕)	-6.233139	106.812124	Korean Restaurant

Based on the data, it is known that in the neighborhood in Jakarta have 234 unique venues categories.

d. Occurrence Frequency Venues Category For Each Neighboorhod In Jakarta

Below is data of grouped neighborhood and by taking the mean of the frequency of occurrence of each category.

	Neighborhoods	Accessories Store	Acehnese Restaurant	African Restaurant	American Restaurant	Arcade	Art Gallery	Art Museum	Asian Restaurant	Athletics & Sports	...	Veterinarian	Video Game Store	Video Store	Vietnamese Restaurant	Water Park	Wine Bar	Wings Joint
0	AIS Indonesia	0.0	0.00	0.0	0.00	0.00	0.02	0.01	0.03	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.01
1	ASEAN MRT station	0.0	0.01	0.0	0.00	0.00	0.00	0.00	0.07	0.01	...	0.0	0.0	0.0	0.0	0.0	0.00	0.01
2	Atma Jaya Catholic University of Indonesia	0.0	0.00	0.0	0.00	0.02	0.01	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
3	BTPN Towers	0.0	0.00	0.0	0.00	0.00	0.01	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
4	Balai Sarbini	0.0	0.00	0.0	0.00	0.00	0.01	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
58	Tebet railway station	0.0	0.00	0.0	0.00	0.00	0.02	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
59	Tebet, South Jakarta	0.0	0.00	0.0	0.00	0.00	0.02	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
60	The Pakubuwono	0.0	0.00	0.0	0.00	0.01	0.00	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.01	0.01
61	The Peak Twin Towers	0.0	0.00	0.0	0.00	0.00	0.01	0.00	0.01	0.00	...	0.0	0.0	0.0	0.0	0.0	0.00	0.00
62	The Plaza Semanggi	0.0	0.00	0.0	0.01	0.00	0.00	0.00	0.02	0.01	...	0.0	0.0	0.0	0.0	0.0	0.01	0.00

e. Filtered Data for "Korean Restaurant" in Jakarta

Below is data of filtered venue category only for Korean Restaurant in Jakarta's Neighborhood as the focus in this project

	Neighborhoods	Korean Restaurant
0	AIS Indonesia	0.00
1	ASEAN MRT station	0.00
2	Atma Jaya Catholic University of Indonesia	0.00
3	BTPN Towers	0.01
4	Balai Sarbini	0.01
...	...	...
58	Tebet railway station	0.02
59	Tebet, South Jakarta	0.02
60	The Pakubuwono	0.02
61	The Peak Twin Towers	0.01
62	The Plaza Semanggi	0.04

f. Clustered Data for Korean Restaurant in Neighborhood in Jakarta

The clustering from k-Means categorized the neighborhoods into three clusters based on the occurrence of "Korean Restaurant" in Jakarta, Indonesia. These are the result of k-Means clustering:

1. Cluster 0: Neighborhoods in Jakarta with low number of Korean Restaurant

]:

	Neighborhood	Korean Restaurant	Cluster Labels	Latitude	Longitude
0	AIS Indonesia	0.0	0	-6.278240	106.828580
28	Lebak Bulus Grab MRT station	0.0	0	-6.289210	106.776030
37	Pancoran, South Jakarta	0.0	0	-6.255480	106.843910
38	Pasar Minggu Baru railway station	0.0	0	-6.171480	106.826490
39	Pasar Minggu railway station	0.0	0	-6.292650	106.827910
41	Pejaten Village	0.0	0	-6.280500	106.829050
43	Pesanggrahan, South Jakarta	0.0	0	-6.236400	106.754470
21	Jeruk Purut Cemetery	0.0	0	-6.279910	106.816500
44	Pondok Indah	0.0	0	-6.268570	106.780560
45	Pondok Indah Mall	0.0	0	-6.266260	106.783350
18	Jakarta Intercultural School	0.0	0	-6.286060	106.782530
17	Jagakarsa	0.0	0	-6.343080	106.817450
46	Ragunan Zoo	0.0	0	-6.301940	106.814380
29	Lenteng Agung railway station	0.0	0	-6.323830	106.839080
15	Golden Triangle of Jakarta	0.0	0	-6.139410	106.809250
13	Fatmawati MRT station	0.0	0	-6.292530	106.793400

2. Cluster 1: Neighborhoods in Jakarta with high number of Korean Restaurant

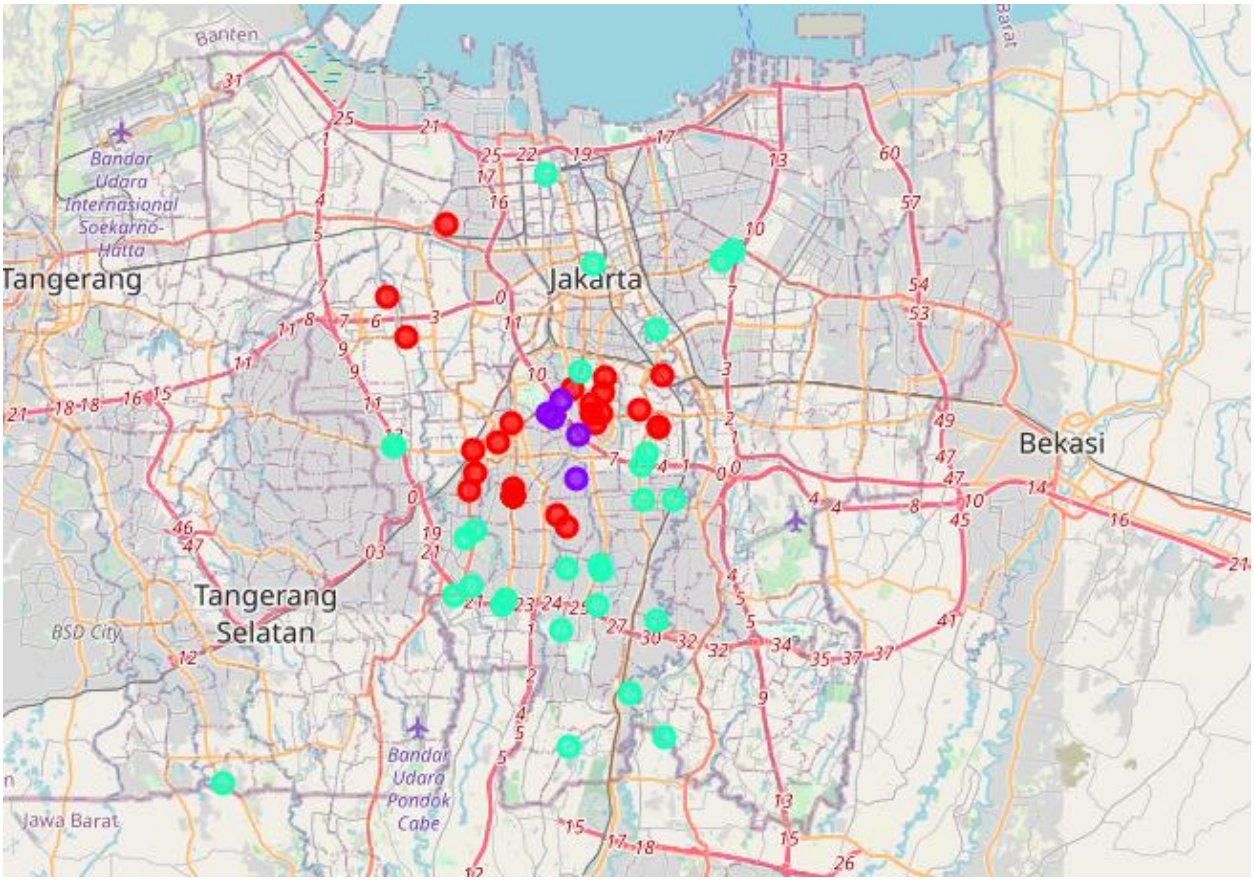
[56]:

	Neighborhood	Korean Restaurant	Cluster Labels	Latitude	Longitude
55	Sudirman Central Business District	0.06	1	-6.22577	106.81183
36	Pacific Place Jakarta	0.06	1	-6.22450	106.80974
62	The Plaza Semanggi	0.04	1	-6.21959	106.81440
7	Capital Place Jakarta	0.05	1	-6.23201	106.82067

3. Cluster 2: Neighborhoods in Jakarta with moderate number of Korean Restaurant

	Neighborhood	Korean Restaurant	Cluster Labels	Latitude	Longitude
30	List of regencies and cities in Jakarta	0.010000	2	-6.253740	106.797310
60	The Pakubuwono	0.020000	2	-6.234881	106.791954
59	Tebet, South Jakarta	0.010000	2	-6.229390	106.849710
58	Tebet railway station	0.010000	2	-6.229390	106.849710
3	BTPN Towers	0.010000	2	-6.253740	106.797310
4	Balai Sarbini	0.010000	2	-6.253740	106.797310
6	Blok M, Jakarta	0.010000	2	-6.157031	106.773199
53	South Jakarta	0.010000	2	-6.253740	106.797310
52	Singapore International School, Indonesia	0.011236	2	-6.197270	106.758760
9	Ciputra World Jakarta	0.010000	2	-6.253740	106.797310
50	Setiabudi Astra MRT station	0.010000	2	-6.209560	106.821590
49	Setiabudi	0.010000	2	-6.221240	106.825280
14	Gandaria City	0.020000	2	-6.245720	106.783930
61	The Peak Twin Towers	0.010000	2	-6.253740	106.797310
47	Senayan City	0.020000	2	-6.228000	106.796760
19	Jalan H.R. Rasuna Said	0.020000	2	-6.211365	106.830520

Below are the visualization of those three category clustering of neighborhoods in Jakarta:



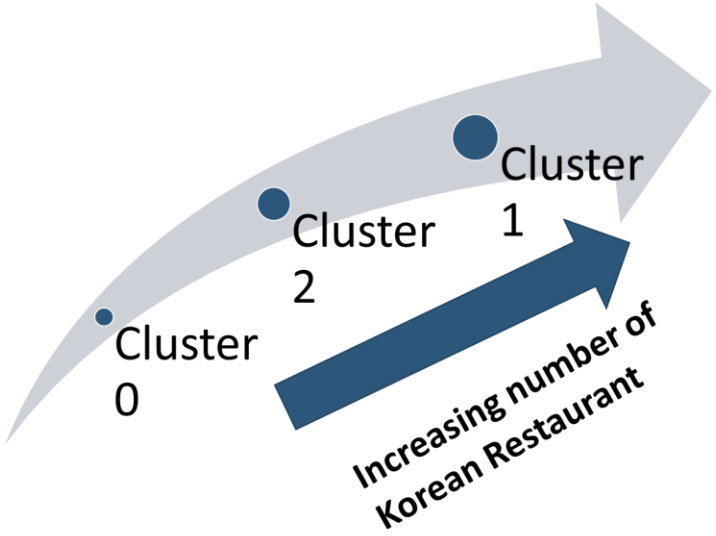
Legends

Red Color: Cluster 0

Purple Color: Cluster 1

Mint Color: Cluster 2

4. DISCUSSION



As observed in the clustered map of Korean Restaurant in the Jakarta neighborhood, it is known that the most Korean Restaurant is concentrated in central of South Jakarta districts which is indicated by Cluster 1(Purple Color). On the other hand, area in Cluster 0 (Red Color) shows the least concentrated number of Korean Restaurant. Meanwhile the Cluster 2 (Mint Color) as located in south area of South Jakarta District

Based on these observation, it could be said that the best location to open new Korean Restaurant in Jakarta would be in the Cluster 0 area. This would be great opportunity to open New Korean Restaurant in this area since it has very little to no competition from existing Korean Restaurant. In addition, the location is also strategic enough as area position is not far from center of the city.

Meanwhile, Korean restaurant in cluster 1 are likely suffering from intense competition due to oversupply and high concentration of Korean restaurant. From another perspective, this also shows that the oversupply of Korean restaurant mostly happened in the central area of the city.

Therefore, this project recommends property developers to capitalize on these findings to open new Korean restaurant in neighborhoods in cluster 0 with little to no competition. Property developers with unique selling propositions to stand out from the competition can also open new Korean restaurant in neighborhoods in cluster 2 with moderate competition. Lastly, property developers are advised to avoid neighborhoods in cluster 1 which already have high concentration of Korean restaurant and suffering from intense competition.

6. LIMITATION AND SUGGESTIONS FOR FUTURE RESEARCH

This project only considered one factor which is the frequency of occurrence of Korean Restaurant in the Neighborhood in Jakarta as the suggestion for property developer to build New Korean Restaurant in Jakarta. Meanwhile there are a lot of other factors that affecting this decision, like population, income, and also demographical data. However, to the best knowledge of this researcher those data are still not available yet to the neighborhood level as required by this project. Therefore, for future research, it would be better if it could devise a methodology to estimate those data to be used in clustering algorithm to determine the preferred location to open New Korean Restaurant in Jakarta.

7. CONCLUSION

Therefore, this project recommends property developers to capitalize on these findings to open new Korean restaurant in neighborhoods in cluster 0 with little to no competition. Property developers with unique selling propositions to stand out from the competition can also open new Korean restaurant in neighborhoods in cluster 2 with moderate competition. Lastly, property developers are advised to avoid neighborhoods in cluster 1 which already have high concentration of Korean restaurant and suffering from intense competition.