Coursera Capstone – The Battles of Neighborhoods

Introduction – Business Problem

Background :

Jakarta is the most developed and populated area with higher salary income in Indonesia, it has lots of place to offers for businesses to grow to a higher scale and level and variety of market.

However, it has a lots of competitor and therefore needs to be reviewed on where companies should put their location with a lot of consideration so they can benefit or take the smallest risk possible if things went south.

Problem :

As one of the newcomer of fast food Industries competitor, FOODINC needs to know where they should put their restaurant at in parts of Jakarta Selatan (they already have branch in other cities). With limited budget at risk, they want to know where is the best place for them with the smallest risk possible.

Target Audience :

By hiring me as a data scientist, weve been conducting a focus group discussion whether what kind of possible indicator might be needed and which place is the best to be recommended to their CEO, I would also expected to give reports in the most layman terms possible for all of the employees of FOODINC to understand.

Success Criteria :

The criteria for the management to consider is its demand and population of neighborhoods.

Data Description :

As we need to explore, segment, and cluster the neighborhoods (kelurahan) in the city of Jakarta Selatan, the neighborhood data is the key for this project. Unfortunately, the data isn't available on the internet in a database or csv format. Hence we need to scrap it from :
:https://selatan.jakarta.go.id/page-statistik-kependudukan (population statistics)

https://www.infojabodetabek.com/kode-pos-di-kota-jakarta-selatan/ (postal codes / neighborhood)

both scraping is then cleaned up and are in structured format of :

1. Postal Codes
2. Borough
3. Neighborhood
4. Latitude
5. Longitude

We can acquire latitude and longitude for each neighborhood by using python library Geopy and run through all neighborhood.

	Kode Pos	Kecamatan	Kelurahan	Latitude	Longitude
0	12430	Cilandak	Cilandak Barat	-6.290266	106.796215
1	12410	Cilandak	Cipete Selatan	-6.272540	106.805768
2	12420	Cilandak	Gandaria Selatan	-6.271562	106.795172
3	12440	Cilandak	Lebak Bulus	-6.302568	106.780354
4	12450	Cilandak	Pondok Labu	-6.309144	106.797938

Table 1. Data of Jakarta Selatan

Data Features :

We will leverage our features using reliable information from a FourSquare API that let us explore various venues in each neighborhood. The information will be :

1. Neighborhood
2. Neighborhood Population
3. Latitude
4. Longitude
5. Venue Name
6. Venue Category
7. Venue Latitude
8. Venue Longitude

We will then cluster these Neighborhood into similiarities and find out which region and place is the best for FOODINC to place their new fast food restaurant.

Result :

Using K-Means Clustering technique , we will segment neighborhood into 4 clusters. With the result apparent in the picture below. These clusters are based on similiarities of top venues and population of each neighborhoods. This cluster are used to find the best place for FOODINC according to the success criteria.

	Borough	Cluster Labels	Man Population	Woman Population	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
0	Cilandak	0.0	15976.0	15829.0	Coffee Shop	Café	Asian Restaurant	Indonesian Restaurant	Pizza Place	Fast Food Restaurant	Ice Cream Shop	Bookstore
1	Cilandak	0.0	13040.0	13162.0	Moving Target	Javanese Restaurant	Arcade	Food Court	Café	Wings Joint	Padangnese Restaurant	Nightclub
3	Cilandak	0.0	29928.0	29878.0	Convenience Store	Indonesian Restaurant	Café	Bubble Tea Shop	Asian Restaurant	Shabu-Shabu Restaurant	Paper / Office Supplies Store	Coffee Shop
9	Jagakarsa	0.0	35441.0	34704.0	College Residence Hall	Light Rail Station	Theme Park	Wings Joint	Padangnese Restaurant	Music Store	Music Venue	Nightclub
11	Kebayoran Baru	0.0	21059.0	20819.0	Post Office	Bakery	Snack Place	Diner	Juice Bar	Karaoke Bar	Nightclub	Noodle House
13	Kebayoran Baru	0.0	5602.0	5661.0	Japanese Restaurant	Coffee Shop	Sushi Restaurant	Hotel	Convenience Store	Restaurant	Pizza Place	Padangnese Restaurant

From the clusters we see that the first cluster offers varieties of restaurants and is near to college residence

it allows FOODINC to gain market shares of due to similiar neighborhoods.

as we see that we need higher demand and lower supplies, and we can see that at borough of Pasar Minggu,at one of its neighborhood it has high population and demand for fast food.

we can also conclude due to college residences in the same cluster, we can make package deals for college students to buy within their pocket money

with that in mind we can conclode that Pasar Minggu is the best place to put a new chain of fast food restaurant!