

The Battles of Neighborhoods

Yulian Adhitama

1. Introduction

1.1. Background

Jakarta (/dʒəˈkɑːrtə/; Indonesian pronunciation: [dʒaˈkarta] (About this soundlisten)), officially the Special Capital Region of Jakarta (Indonesian: Daerah Khusus Ibukota Jakarta), is the capital and a province of Indonesia. Situated on the northwest coast of the world's most populous island of Java, it is the centre of economy, culture and politics of Indonesia with a population of more than ten million as of 2014. Although Jakarta only covers 699.5 square kilometres (270.1 sq mi), the smallest among any Indonesian provinces, its metropolitan area covers 6,392 square kilometres (2,468 sq mi); it is the world's second-most populous urban area after Tokyo, with a population of about 30 million as of 2010. Jakarta's business opportunities, as well as its potential to offer a higher standard of living, have attracted migrants from across the Indonesian archipelago, making it a melting pot of numerous cultures. Jakarta is nicknamed the "Big Durian", the thorny strongly-odored fruit native to the region, as the city is seen as the Indonesian equivalent of New York (Big Apple).

1.2. Problem

Jakarta is the largest city and capital of Indonesia. Jakarta lies on the northwest coast of Java at the mouth of the Ciliwung (Liwung River), on Jakarta Bay (an embayment of the Java Sea). It is coextensive with the metropolitan district of Greater Jakarta (Jakarta Raya) and nearly coextensive with the daerah khusus ibukota (special capital district) of Jakarta—the latter also including a number of small offshore islands in the Java Sea. With a large population and as the largest city in Indonesia, opening a new venue could be a great idea. But, the problem is each borough and neighborhood has its own culture and uniqueness. And the question is what is best venue to open in Jakarta? This answer could help anyone who wants to open a new venue in Jakarta.

2. Data acquisition and cleaning

2.1. Data resources

Data resources obtained from Wikipedia https://id.wikipedia.org/wiki/Daftar_kecamatan_dan_kelurahan_di_Daerah_Khusus_Ibukota_Jakarta. This data provide 6 borough and 43 neighborhoods in total.

2.2. Data Cleaning

Data read from resources using pandas built in function `read_html`. After that, the neighborhoods that have borough name of 'Kepulauan Seribu' dropped, since that borough is not business borough with large population. Also, only neighborhood and borough columns kept, since we only interested in both information only. Next, geopy library is used to extract latitude and longitude of given neighborhoods and then the data showed in map form with folium library.

3. Methodology

3.1. Show Jakarta Map

I choose to show the map of Jakarta as shown below:

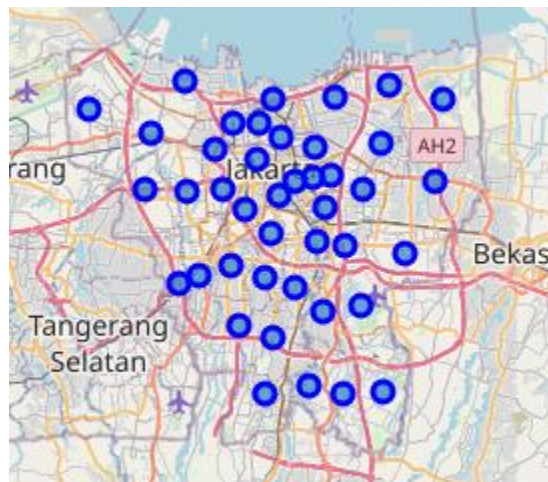


Figure 1. Map of Jakarta

3.2. Get nearby venues

Next we define function to find all the nearby venues of given neighborhoods. Once we created the function, then we pass dataframe that contains neighborhood names along with latitudes and longitudes to get nearby venues.

	Neighborhood	Borough	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Cempaka Putih	Jakarta Pusat	-6.181214	106.868548	Mie Aceh Bungong Cempaka	-6.180464	106.871099	Acehnese Restaurant
1	Cempaka Putih	Jakarta Pusat	-6.181214	106.868548	Bika Ambon Medan Sylvie	-6.181566	106.866629	Bakery
2	Cempaka Putih	Jakarta Pusat	-6.181214	106.868548	Pizza Hut	-6.181644	106.872073	Pizza Place
3	Cempaka Putih	Jakarta Pusat	-6.181214	106.868548	Pizza Hut	-6.179067	106.866189	Pizza Place
4	Cempaka Putih	Jakarta Pusat	-6.181214	106.868548	Arcici Swiming Pool™	-6.184377	106.867813	Pool

Figure 2. Each neighborhoods along with nearby venues

3.3. Convert Categorical Variable into Numerical Variable

Next, we convert Venue columns, which is categorical variable type into numerical variable using one hot encoding. Then we group each venue by its mean based on the given neighborhood.

Neighborhood	Yoga Studio	Accessories Store	Acehnese Restaurant	Airport Terminal	American Restaurant	Arcade	Art Gallery	Art Museum	Asian Restaurant	...	Supermarket	Sushi Restaurant	Thai Restaurant
0	Cakung	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
1	Cempaka Putih	0.00	0.000000	0.142857	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
2	Cengkareng	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
3	Cilandak	0.00	0.000000	0.000000	0.000000	0.000000	0.250000	0.000000	0.000000	...	0.000000	0.000000	0.000000
4	Cilincing	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
5	Cipayung	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
6	Ciracas	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
7	Duren Sawit	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
8	Gambir	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
9	Grogol Petamburan	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.111111	...	0.000000	0.027778	0.000000
10	Jagakarsa	0.00	0.000000	0.200000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
11	Jatinegara	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
12	Johar Baru	0.00	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000	0.000000	...	0.000000	0.000000	0.000000
13	Kalideres	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000
14	Kebayoran Baru	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.027027	0.081081	0.000000
15	Kebayoran Lama	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000

Figure 3. Each mean venues based on neighborhood

3.4. Find the Top 10 Most Common Venues

The next thing after find each venue mean based on neighborhood is to sort them and find the top 10 most common venues.

	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	Cakung	Gas Station	Lounge	Soup Place	Food & Drink Shop	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner	Vietnamese Restaurant
1	Cempaka Putih	Pizza Place	Acehnese Restaurant	Indonesian Meatball Place	Pool	Bakery	BBQ Joint	Vietnamese Restaurant	Diner	Farmers Market	Electronics Store
2	Cengkareng	Coffee Shop	Bubble Tea Shop	Food Truck	Cupcake Shop	Dance Studio	Deli / Bodega	Department Store	Dessert Shop	Food & Drink Shop	Convenience Store
3	Cilandak	Food Truck	Health & Beauty Service	Soup Place	Art Gallery	Diner	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop
4	Cilincing	Park	Shopping Mall	Diner	Vietnamese Restaurant	Dim Sum Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Dessert Shop

Figure 4. Top 10 Most Common Venues each neighborhood

3.5. Cluster the Neighborhood Based on Venues

For cluster the neighborhood based on venues, I use K-Means Clustering algorithm. I choose $k=5$ to cluster the neighborhood into 5 cluster, and I choose to initialize centroid at random point. Library used is Scikit Learn. This is the final method to get the final result. The result of this method is explained in the next section, in chapter 4: result.

4. Result

4.1. Clusters Map

The map of cluster neighborhood is shown below:

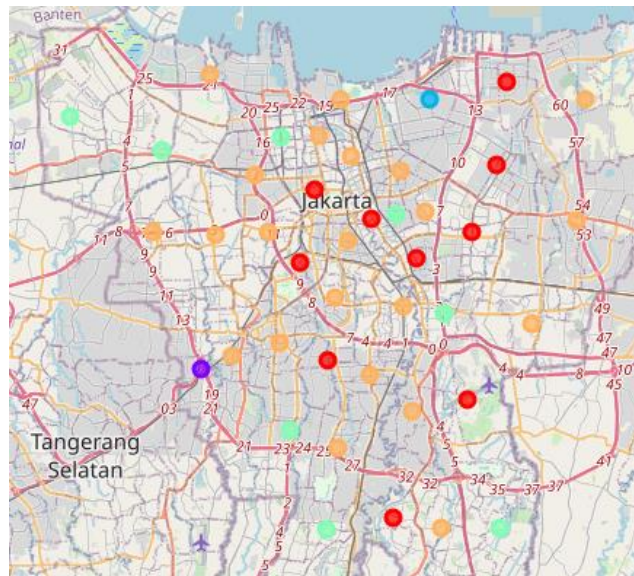


Figure 5. Cluster Map

From the above map, we know that each neighborhood is represented in point with different color, where same color represent same cluster. Here, we see that we have 5 different color, red, orange, green, blue, and purple, each represent different clusters. Red for cluster 1, purple for cluster 2, blue for cluster 3, green for cluster 4, and orange for cluster 5.

4.2. Explore Each Cluster

4.2.1. Cluster 1

	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
1	Gambir	Indonesian Restaurant	Seafood Restaurant	Café	History Museum	Hotel	Clothing Store	Office	Food Truck	Breakfast Spot	Bakery
6	Senen	Hotel	Indonesian Restaurant	University	History Museum	Dance Studio	Noodle House	Art Gallery	Donut Shop	Vietnamese Restaurant	Diner
7	Tanah Abang	Indonesian Restaurant	Coffee Shop	Seafood Restaurant	Soup Place	Noodle House	Food Truck	Javanese Restaurant	Japanese Restaurant	Fruit & Vegetable Store	Pizza Place
9	Kelapa Gading	Indonesian Restaurant	Asian Restaurant	Chinese Restaurant	Japanese Restaurant	Korean Restaurant	Steakhouse	Seafood Restaurant	Javanese Restaurant	Food Court	Coffee Shop
10	Koja	Indonesian Restaurant	Pizza Place	Arcade	Vietnamese Restaurant	Dim Sum Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner
20	Makasar	Indonesian Restaurant	Concert Hall	Airport Terminal	Chinese Restaurant	Asian Restaurant	Deli / Bodega	Dance Studio	Cupcake Shop	Department Store	Food Court
21	Matraman	Coffee Shop	Snack Place	Pizza Place	Indonesian Restaurant	Dim Sum Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner
22	Pasar Rebo	Yoga Studio	Bakery	Noodle House	Indonesian Restaurant	Diner	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop
23	Pulo Gadung	Indonesian Restaurant	Convenience Store	Athletics & Sports	Fast Food Restaurant	Pizza Place	Café	Hardware Store	Sandwich Place	Breakfast Spot	Bubble Tea Shop
28	Mampang Prapatan	Noodle House	Indonesian Restaurant	Fast Food Restaurant	Bakery	Vietnamese Restaurant	Diner	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop

Figure 6. Result of Cluster 1

From above table, we see that the 1st most common venue is Indonesian Restaurant. Thus, we can conclude that cluster 1 is belong to Indonesian Restaurant. And the best venue to open in this cluster is Indonesian Restaurant, which represented in red point in map.

4.2.2. Cluster 2

	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
31	Pesanggrahan	Pizza Place	Dim Sum Restaurant	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner	Dessert Shop	College Academic Building

Figure 7. Result of Cluster 2

From above table, we see that the 1st most common venue is Pizza Place. Thus, we can conclude that cluster 2 is belong to Pizza Place. And the best venue to open in this cluster is Pizza Place, which represented in purple point in map.

4.2.3. Cluster 3

	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
13	Tanjung Priok	Bakery	BBQ Joint	Vietnamese Restaurant	Food Truck	Food & Drink Shop	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop

Figure 8. Result of Cluster 3

From above table, we see that the 1st most common venue is Bakery. Thus, we can conclude that cluster 3 is belong to Bakery. And the best venue to open in this cluster is Bakery, which represented in blue point in map.

4.2.4. Cluster 4

	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
2	Johar Baru	Food Truck	Arcade	Café	Department Store	Food & Drink Shop	Convenience Store	Gym / Fitness Center	Gym	Cupcake Shop	Dance Studio
15	Cipayung	Food Truck	Food Court	Convenience Store	Cupcake Shop	Dance Studio	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner
18	Jatinegara	Garden	Satay Restaurant	Food Truck	Gym	Italian Restaurant	Vietnamese Restaurant	Electronics Store	Dumpling Restaurant	Donut Shop	Diner
24	Cilandak	Food Truck	Health & Beauty Service	Soup Place	Art Gallery	Diner	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop
25	Jagakarsa	Pharmacy	Department Store	Convenience Store	Food Truck	Acehnese Restaurant	Health & Beauty Service	Hardware Store	History Museum	Cupcake Shop	Dance Studio
34	Cengkareng	Coffee Shop	Bubble Tea Shop	Food Truck	Cupcake Shop	Dance Studio	Deli / Bodega	Department Store	Dessert Shop	Food & Drink Shop	Convenience Store
37	Tambora	Food Truck	Garden	Convenience Store	Train Station	Department Store	Indonesian Restaurant	Miscellaneous Shop	Farmers Market	Electronics Store	Dumpling Restaurant
39	Kalideres	Ski Area	Convenience Store	Bookstore	Food Truck	Soccer Field	Fried Chicken Joint	Indonesian Restaurant	Donut Shop	Farmers Market	Electronics Store

Figure 9. Result of Cluster 4

From above table, we see that the 1st most common venue is Food Truck. Thus, we can conclude that cluster 4 is belong to Food Truck. And the best venue to open in this cluster is Food Truck, which represented in green point in map.

4.2.5. Cluster 5

	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	Cempaka Putih	Pizza Place	Acehnese Restaurant	Indonesian Meatball Place	Pool	Bakery	BBQ Joint	Vietnamese Restaurant	Diner	Farmers Market	Electronics Store
1	Kemayoran	Noodle House	Arcade	Spa	Vietnamese Restaurant	Dim Sum Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner
2	Menteng	Food Truck	Indonesian Restaurant	Coffee Shop	Park	Deli / Bodega	Department Store	Miscellaneous Shop	Dessert Shop	Massage Studio	Electronics Store
3	Sawah Besar	Indonesian Restaurant	Convenience Store	Noodle House	Karaoke Bar	Asian Restaurant	Pet Store	Diner	Automotive Shop	Camera Store	Chinese Restaurant
4	Cilincing	Park	Shopping Mall	Diner	Vietnamese Restaurant	Dim Sum Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Dessert Shop
5	Pademangan	Hotel	Food Truck	Theme Park	Shopping Mall	Bowling Alley	Athletics & Sports	Asian Restaurant	Seafood Restaurant	Department Store	Dessert Shop
6	Penjaringan	Pharmacy	Food Truck	Theme Park	Indonesian Restaurant	Restaurant	Dessert Shop	Dim Sum Restaurant	Electronics Store	Dumpling Restaurant	Donut Shop
7	Cakung	Gas Station	Lounge	Soup Place	Food & Drink Shop	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Diner	Vietnamese Restaurant
8	Ciracas	Coffee Shop	Metro Station	Diner	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Dim Sum Restaurant	Concert Hall
9	Duren Sawit	Indonesian Meatball Place	Coffee Shop	Mediterranean Restaurant	Convenience Store	Cupcake Shop	Dance Studio	Deli / Bodega	Department Store	Food & Drink Shop	Dim Sum Restaurant
10	Kramat Jati	Diner	Chinese Restaurant	Arcade	Vietnamese Restaurant	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop	Dessert Shop

Figure 10. Result of Cluster 5

11	Kebayoran Baru	Japanese Restaurant	Coffee Shop	Hotel	Sushi Restaurant	Korean Restaurant	Cupcake Shop	Fried Chicken Joint	Restaurant	Lounge	Bubble Tea Shop
12	Kebayoran Lama	Bakery	Chinese Restaurant	Javanese Restaurant	Pharmacy	Pizza Place	Diner	Bridal Shop	Restaurant	Donut Shop	Seafood Restaurant
13	Pancoran	Noodle House	Asian Restaurant	Javanese Restaurant	Bakery	Music Venue	Fast Food Restaurant	Farmers Market	Electronics Store	Dumpling Restaurant	Donut Shop
14	Pasar Minggu	Asian Restaurant	Café	Food Court	Bus Line	Convenience Store	Department Store	Deli / Bodega	Dance Studio	Dessert Shop	Cupcake Shop
15	Setiabudi	Coffee Shop	Hotel	Asian Restaurant	Food Court	Japanese Restaurant	Steakhouse	Multiplex	Performing Arts Venue	Shopping Mall	Massage Studio
16	Tebet	Indonesian Restaurant	Café	Bakery	Asian Restaurant	Karaoke Bar	Dessert Shop	Convenience Store	Restaurant	Indonesian Meatball Place	Sushi Restaurant
17	Grogol Petamburan	Noodle House	Chinese Restaurant	Asian Restaurant	Indonesian Restaurant	Hotel	Pizza Place	Coffee Shop	Fast Food Restaurant	Bookstore	Seafood Restaurant
18	Taman Sari	Chinese Restaurant	Noodle House	Asian Restaurant	Hotel	Convenience Store	Karaoke Bar	Coffee Shop	Steakhouse	Restaurant	Indonesian Restaurant
19	Kebon Jeruk	Noodle House	Concert Hall	Asian Restaurant	Café	Food Truck	Juice Bar	Bakery	Indonesian Restaurant	Coffee Shop	Auto Dealership
20	Palmerah	Coffee Shop	Pizza Place	Noodle House	Asian Restaurant	Multiplex	Restaurant	Medical Center	Convenience Store	Gym	Hotel
21	Kembangan	Asian Restaurant	Japanese Restaurant	Chinese Restaurant	Coffee Shop	Café	Clothing Store	Food Court	Korean Restaurant	Pizza Place	Department Store

Figure 11. Result of Cluster 5 (continued)

From above table, we see that the 1st most common venues is Japanese Restaurant, Asian Restaurant, Noodle House, and Chinese Restaurant, which we can cluster them into Asian Food Restaurant. Thus, we can conclude that cluster 5 is belong Asian Food Restaurant. And the best venue to open in this cluster is Asian Food Restaurant, which represented in orange point in map.

5. Discussion

From the result, we can see that the cluster divided into 5 different groups. Red for cluster 1, purple for cluster 2, blue for cluster 3, green for cluster 4, and orange for cluster 5. The neighborhood is not equally separated into 5 different clusters. Cluster 1, 4, and 5 have many neighborhoods, whereas cluster 2 and 3 only have 1 neighborhood. We can increase the number of neighborhood of these clusters by increasing radius of nearby venues. Or, if we think we should not increase the radius of nearby venues each neighborhood, we can drop these 2 clusters that have 1 neighborhood by setting $k=3$. So that 1 neighborhood that previously assigned to each of cluster 2 and 3, now assigned to either cluster 1, 4, or 5.

6. Conclusion

In this study, we cluster each neighborhood in Jakarta based on most common venue in each neighborhood with help of foursquare API. We cluster into 5 different group using K-Means Clustering algorithm. The result is we get 5 different cluster which plot into 5 different color in map using folium library. Red for cluster 1, purple for cluster 2, blue for cluster 3, green for cluster 4, and orange for cluster 5. The 1st most common venue from cluster 1 to cluster 5 is food place, like restaurant, food truck, or bakery. So, we conclude that food place is a good idea if we want to open new venue in Jakarta. And based on the result, here is the recommendation of what kind of food place if we want to open in each cluster:

1. Cluster 1: Indonesian Restaurant
2. Cluster 2: Pizza Place
3. Cluster 3: Bakery
4. Cluster 4: Food Truck
5. Cluster 5: Asian Food Restaurant