Jakarta : Battle of Neighborhood

Business Problem

Why Jakarta?

- The second most populous urban area in the world after Tokyo and 20% of Indonesia's urban population.
  - The capital of Indonesia and the largest city in Indonesia
- One of the fastest growing economies in the world.
  - Reported as the highest return on investment for luxury real estate in 2014 compared to any other city on earth

Why does it matter?

- Many investors are eager to open their hotel business in Jakarta nowadays, which makes it so important that they choose the right location.
- However…
  - It is not easy to get and process the information
  - It requires an analytical thinking to determine the right location

Data Definition

Center of Jakarta is defined as ‘Setiabudi’ area and its coordinate along with neighborhoods will be obtained using **Google Maps API**.

Number of existing hotels in the neighborhood. The data will be generated using **Foursquare API**.

Number of populations in the neighborhood. For this case, the data will be gathered from **Jakarta Government public data repositories** and converted to csv.

All the required data will be processed and presented in map visualization which allows us to see the location in a quick glance.

Methodology

Data Collection

Define Jakarta city center as Setiabudi area and get coordinates using Google Maps API

Collect the location of hotels within ~6 km using Foursquare

Gather population data from Jakarta public datasets

Exploratory Data Analysis

Calculate the density across neighborhood using heatmaps to identify recommended area with criteria as follow:

- Close to city center with 2.5 km radius
- No existing hotels within 1 km
- High number of populations (more than 150,000) in the district

Clustering

Analyze data using k-means clustering and produce 5 location recommendation.

Data Collection

1

A. Get Neighborhood Candidates using Google Maps API

The surrounding neighborhood is defined as ~6 km away from city center, as circular areas with a radius of 300 meters and its center will be 600 meters apart from each other.

- There are 244 neighborhood candidate centers.

2

B. Get Venue in Neighborhood using Foursquare API

- Collect top 100 venue within 500 meters
- Filter to include only Hotel category
- Calculate venue X, Y
- Standardize district name
- There are 304 venues returned

3

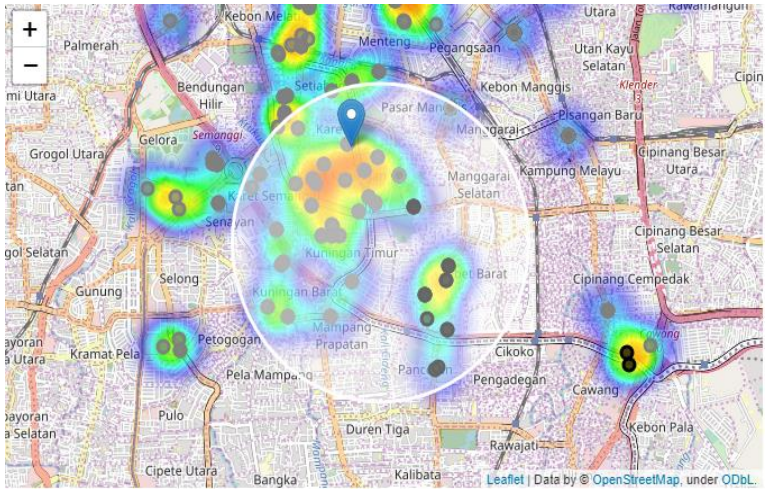
C. Get Population Data

- Merge by district name
- Define high population as more than 150,000

Data Collection – Final Dataset

Neighborhood	Neighborhood_Latitude	Neighborhood_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category	Kecamatan	Venue_x	Venue_y	Nama_Kecamatan	Pria	Wanita	Total_Population	Is_PopulationHigh
5, Jl. Iskandarsyah I No.18, RT.5/RW.4, Melawa...	-6.245414	106.805719	Hotel GranDhika Iskandarsyah Jakarta	-6.244813	106.803989	Hotel	Kby. Baru	699587.931649	-690608.756764	Kby. Baru	74982.0	73362.0	148344.0	False
5, Jl. Iskandarsyah I No.18, RT.5/RW.4, Melawa...	-6.245414	106.805719	Hotel Ambhara	-6.243266	106.803678	Hotel	Kby. Baru	699554.124192	-690437.517698	Kby. Baru	74982.0	73362.0	148344.0	False
5, Jl. Iskandarsyah I No.18, RT.5/RW.4, Melawa...	-6.245414	106.805719	favehotel Melawai	-6.244321	106.801337	Hotel	Kby. Baru	699294.688423	-690553.296449	Kby. Baru	74982.0	73362.0	148344.0	False
Jl. Raya Pasar Minggu No.2 B, RT.2/RW.2, Panco...	-6.245282	106.843668	Cipta Hotel Pancoran	-6.247249	106.843335	Hotel	Pancoran	703941.588910	-690893.251751	Pancoran	78803.0	76393.0	155196.0	True
Jl. Raya Pasar Minggu No.2 B, RT.2/RW.2, Panco...	-6.245282	106.843668	Amaris Hotel Pancoran	-6.247507	106.842834	Hotel	Pancoran	703886.021918	-690921.551012	Pancoran	78803.0	76393.0	155196.0	True

Exploratory Data Analysis



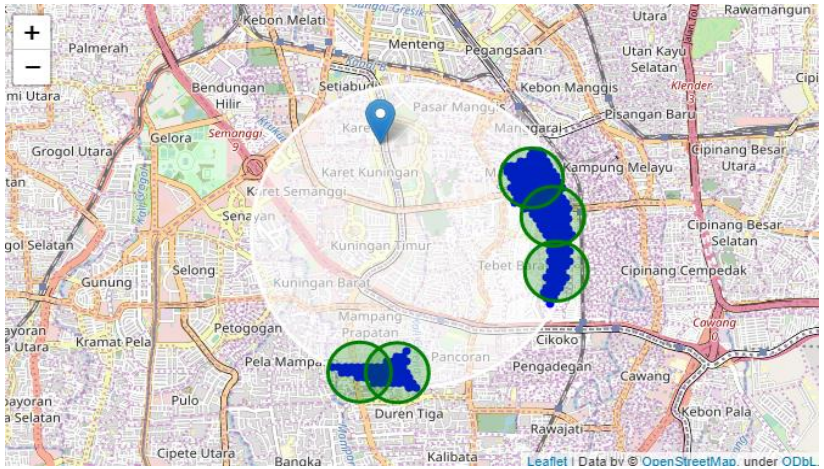
We can focus on the low-density hotel (green heatmap) and high population area (black marker) then generate the new neighborhood with count of hotels for each location.

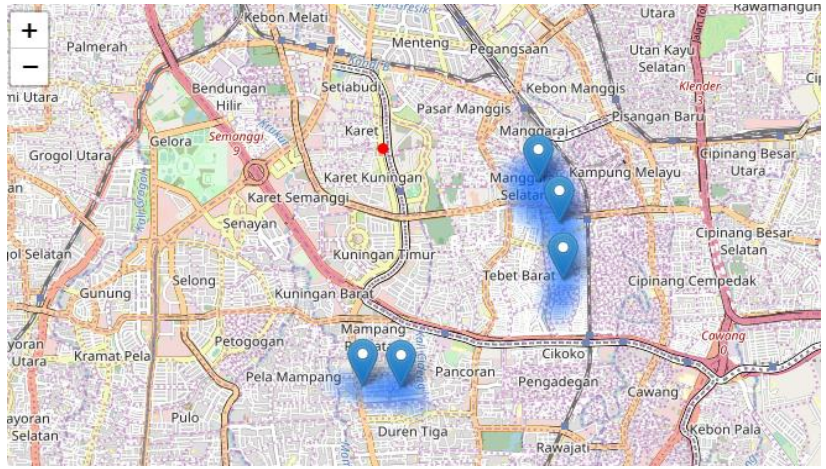
2261 location candidates generated.

Filter the candidates only for location with no hotels within 1 km radius.

199 ‘good’ location candidates generated and ready to be clustered.

Clustering





Results are divided into 5 clusters and the location of centers can be generated using Google Maps API

1. 6, Jl. Tebet Utara III B No.15, RT.7/RW.2, Tebet Tim., Kec. Tebet **=> 2.9km from Setiabudi**
2. Jl. Tegal Parang Selatan V No.47, RT.4/RW.7, Tegal Parang, Kec. Mampang Prpt **=> 3.8km from Setiabudi**
3. Jl. Tebet Timur II G No.2, RT.9/RW.5, Tebet Tim., Kec. Tebet **=> 3.5km from Setiabudi**
4. Jl. Mampang Prpt. Raya No.3a, RT.7/RW.1, Tegal Parang, Kec. Mampang Prpt **=> 3.7km from Setiabudi**
5. 14, Jl. Barkah III No.29, RT.14/RW.2, Manggarai Sel., Kec. Tebet **=> 2.4km from Setiabudi**

Result and Discussion

From analysis that I have done, we can see there's about 144 hotels within 3 km from Jakarta City Center (Setiabudi area) with the closest low-density area detected in south. Additionally, high population density also mainly located in the south so we can focus our interest to the area.

After shifting our focus, I filtered the location candidates to include only area with no hotels within 1 km. Those location were then clustered to create 5 zones of interest which contain greatest number of location candidates. Those are the recommended location for investors to open a new hotel. Finally, I generated the address of the zone centers using reverse geocoding to be used as markers for more detailed analysis based on other factors.

Although we have the recommendation, there are some other factors need to be considered further, such as tourist attraction, socioeconomics and government regulation that are not covered in this analysis. Thus, the result can not be solely used for making decision.

Conclusion

- Problem addressed in this project is how to get the required information and do analysis for determining the best location candidates in opening new hotel by stakeholder/investors.

- By generating datasets and clustering the locations, we have identified 5 areas to be considered for further analysis. The recommendation can be used by investors in decision making process.