COURSERA- IBM DATA SCIENCE CAPSTONE PROJECT

Author : Ramadhani P. Putra

BACKGROUND

- The client wants to build an Indonesian restaurant in Stockholm. He wants to build the restaurant close to the body of water, as he believes it would fulfill his childhood dream.
- He is unaware of the distribution of other similar restaurants as they would be potential competitors.
- He is very well funded and connected to the culinary business in Stockholm, so that supply distribution wouldn't be a major problem.

1. Find a list of neighborhoods in Stockholm city.

2. Find the intensity of certain venues in each neighborhood (*i.e. the intensity of ice cream shop locations in the Stadshagen district*) that potentially become competitors to the client's restaurant.

OBJECTIVES

Yes or No ?

3. Using k-clustering, find the neighborhood where potentially competitor venues has the least value, as this would show some potential site for the restaurant.

4. Cross-correlate the potential site with the surrounding body of water (river, lake, sea).

OBJECTIVES

5. Ultimately, find five neighborhoods in Stockholm that is the most suitable to build the restaurant. Recommend the neighborhood to client with its pros and cons.

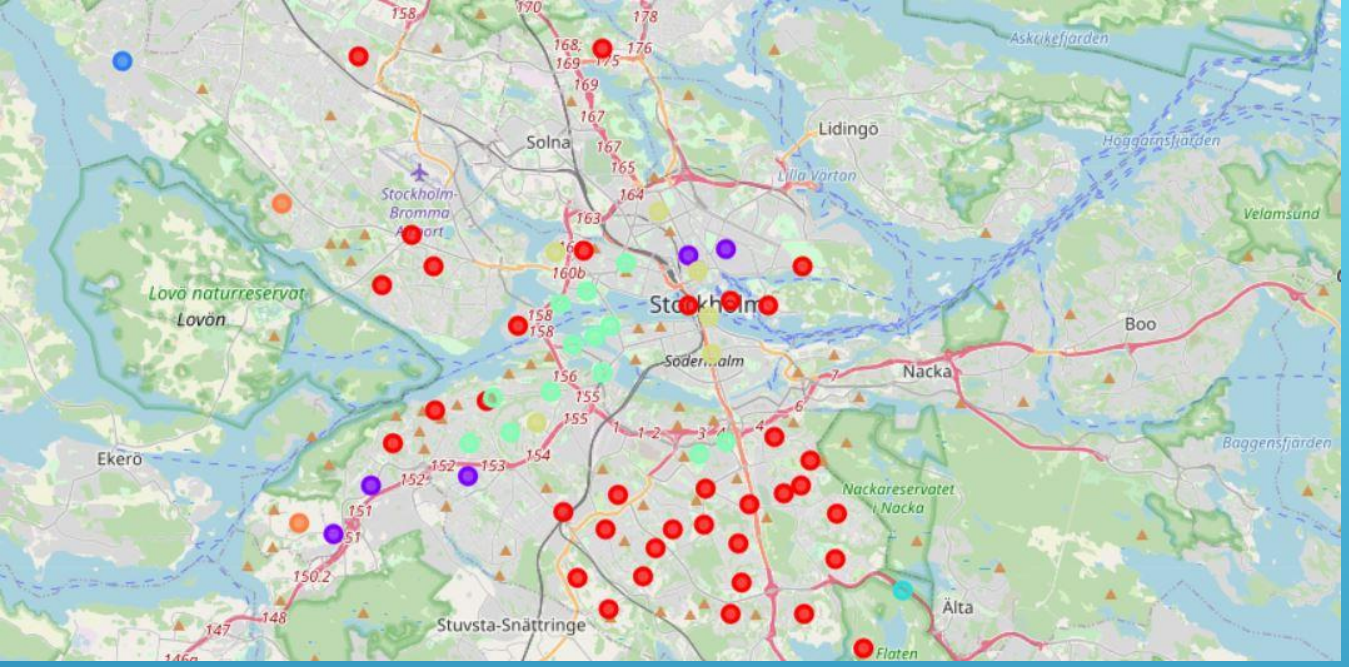
OBJECTIVES

DATA REQUIREMENTS

- The list of Stockholm districts (can be obtained via Wikipedia)
- A digital map of Stockholm, which shows main streets and land allocations.
- Latitude and longitude coordinate data from Geocoder library to assign each neighborhood automatically into the map,

1. Obtain the list of Stockholm neighborhood districts and their coordinates.
2. Superimpose the district location points into the digital map of Stockholm.
3. Using Foursquare API, collect the venues data for each neighborhood. For certain type of venue, calculate its occurence intensity for each neighborhood. We would calculate the occurences of Indonesian restaurant and Asian restaurant.
4. Utilize k-clustering to group and segment Stockholm neighborhoods based on venue intensity.
5. Cross corelate between different venue types, body of water to see which neighborhood can be recommended to the client.
6. Since we can only recommend five neighborhood, we can use other relevant factors to filter and rank the recommended neighborhood.

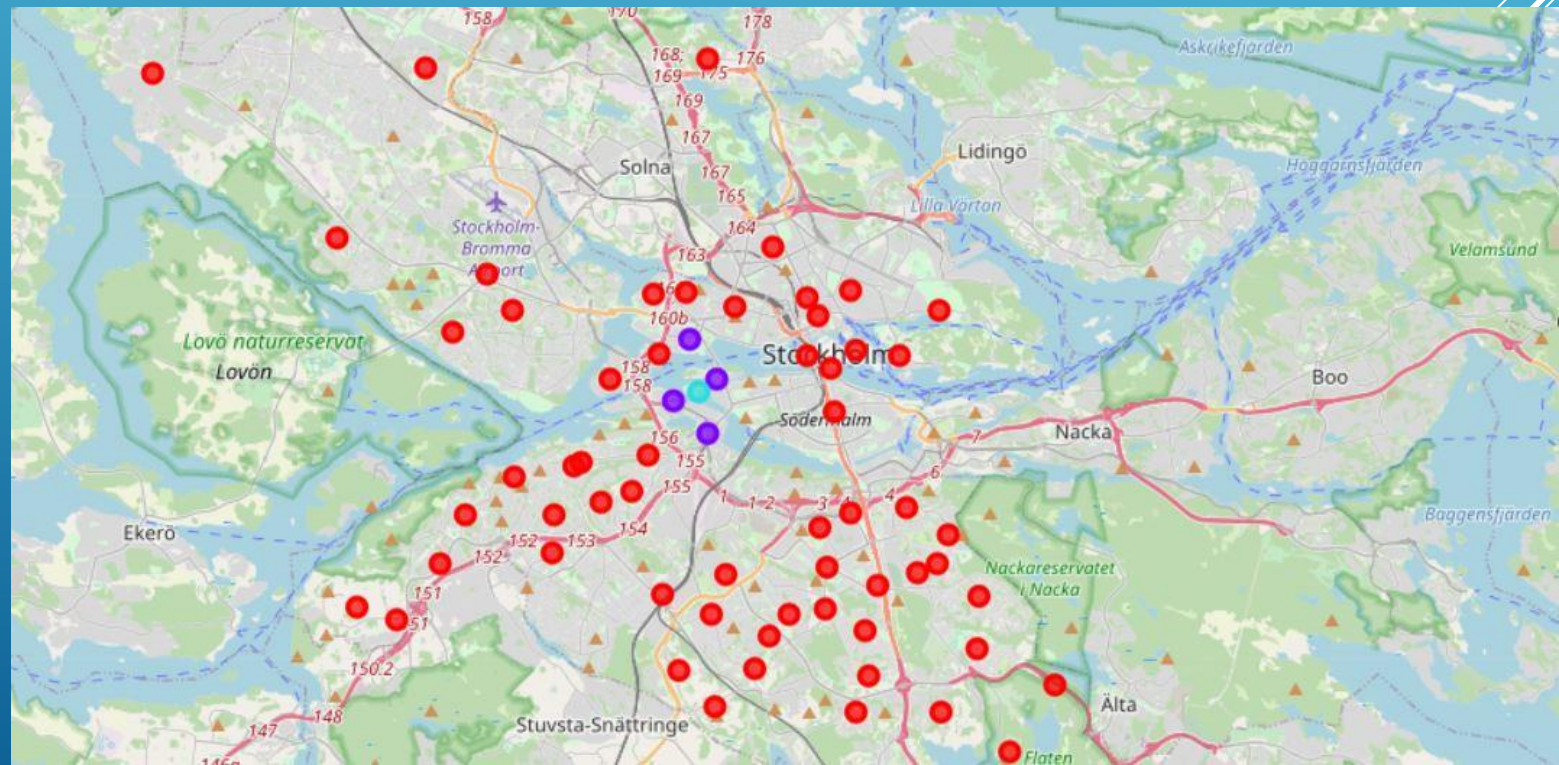
METHODOLOGY



Clustering of Asian restaurants intensities in Stockholm neighborhood.

Clustering of Indonesian restaurants intensities in Stockholm neighborhood.

Because Asian restaurants exist more widely and more intensely than Indonesian restaurant. We would use Asian restaurants clustering as reference later.



RESULTS

1. Areas in easter central Stockholm (except Skeppsholmen) is virtually impossible, since the available land is very limited and is already reserved by government institution/large corporate.
2. Skeppsholmen: This is among the best area because it's close to the city center. Expect very expensive property rent/construction.
3. Farsta, Farsta Strand and Skondal: These areas are close to body of water and somewhat close to the city center.
4. Orhem and Skrubba: These places are not fully developed and difficult to reach for general public. We cannot recommend these areas.
5. Mälarhöjden : This area is close to body of water and somewhat close to the city center via public transport.

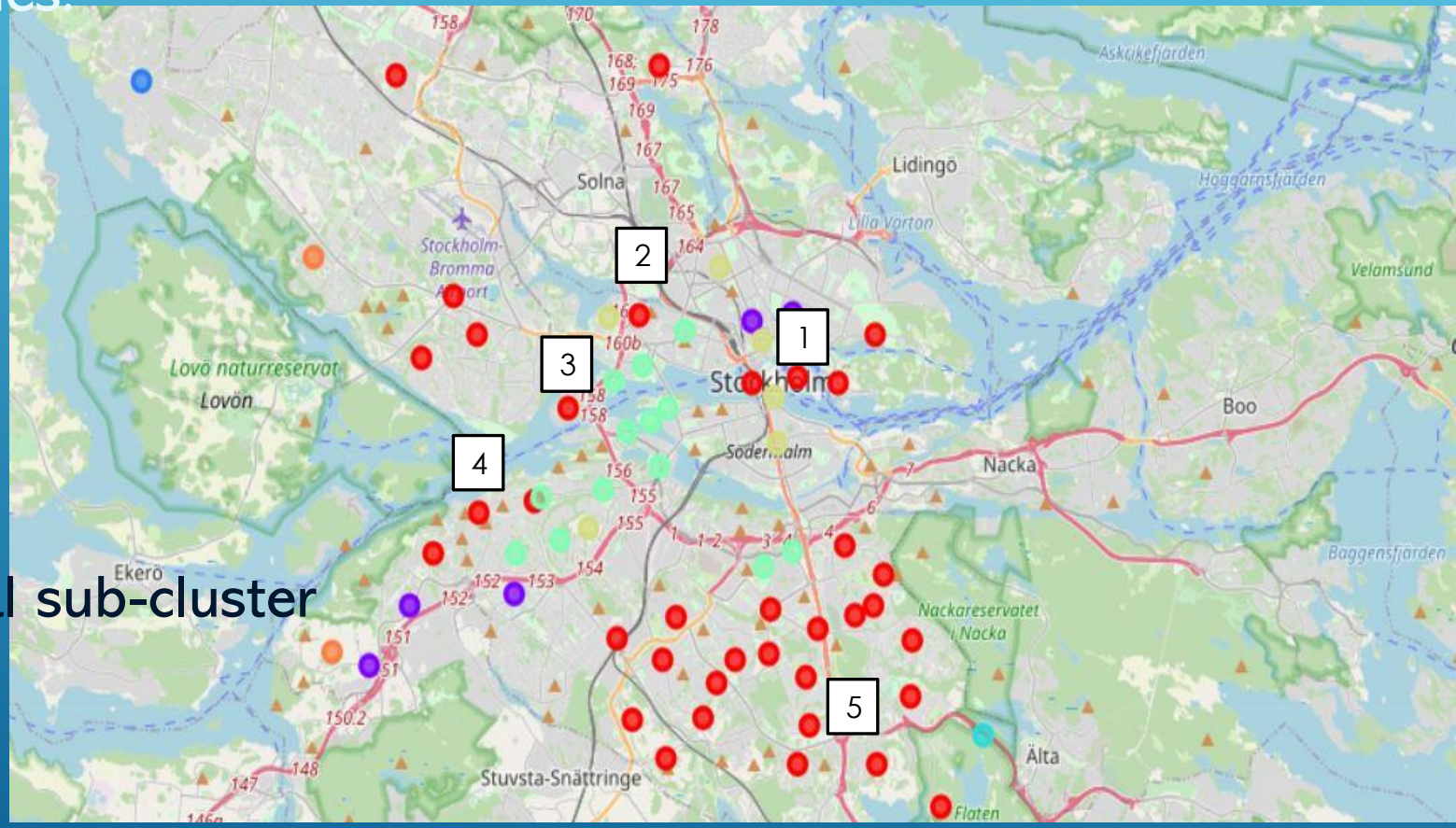
DISCUSSION

6. Bredäng: The land beside the body of water in Bredang is reserved for nature reservation.
7. Axelberg : Good place, but its neighboring area has a few Asian restaurants.
8. Stora Essingen : These areas are close to body of water and somewhat close to the city center.
9. Nockeby : These areas are close to body of water, but quite distant to the city center.
10. Kungshamra : These areas are close to body of water, but quite distant to the city center.
11. Stadshagen :This is among the best area because it's close to the city center. Expect very expensive property rent/construction.

DISCUSSION

From the result, we conclude that we can only recommend a few neighborhood to our client. Below, we ranked these neighborhoods based on multiple factors, primarily due to competitor presence, surrounding body of water, distance to city's center, and other relevant issues:

1. Skeppsholmen
2. Stadshagen
3. Stora Essingen
4. Mälarhöjden
5. Farsta, Farsta Strand and Skondal sub-cluster



CONCLUSION

THE END