PREDICTING THE POTENTIAL LOCATION FOR AN ASIAN SUPERMARKET IN STUTTGART

THI KIM NHUNG PHAN

March 26, 2020

1. Introduction

1.1 Background

Stuttgart is the capital and the largest city of state Baden-Wüttermberg as well as the sixth largest city in Germany. With nice location covered around Neckar river, closed to Black forest and with many companies located in the city, Stuttgart attracts a very large population settle down here, namely 2.8 million people lived in urban area and 5.3 million people in its metropolitan area. A large population with a diversity of nationalities makes Stuttgart an interesting market for retail.

1.2 Problem

However, competition is also fierce in the such high-profit market like Stuttgart. Therefore, analysing the location of competitors as well as of potential customers are crucial in choosing a good place to locate a supermarket. The aim of this project is to find potential locations for an Asian supermarket based on the venues available in Stuttgart.

1.3 Interest

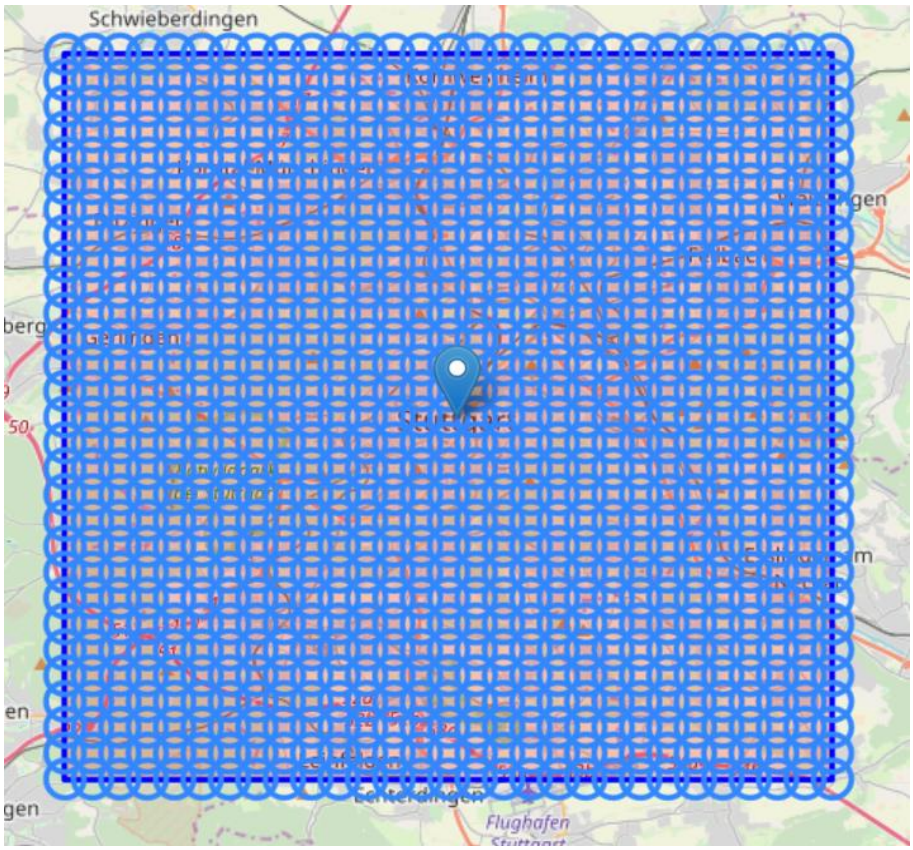
Businessman who would like to open an Asian supermarket in Stuttgart would be in interest. In addition, anyone who has an intention of opening a retail or service business would be interested as well to the approach of this project.

2. Data Description and Preparing

2.1 Data Description

The center's coordination of Stuttgart will be requested using geocoder. From this coordination, a bounding box covered around Stuttgart will be defined and be a base

to generate spaced grid of locations, which are considered as the neighborhoods of Stuttgart.



The centers' coordinations of these neighborhoods will be determined and used to request the venues of its area through Foursquare API (see table below)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	0	48.692019	9.038601	Spatzensee	48.694011	9.043653	Playground
1	0	48.692019	9.038601	café im krankenhaus	48.690379	9.032925	Café
2	1	48.692019	9.048502	Initiativpark	48.692467	9.052749	Athletics & Sports
3	1	48.692019	9.048502	Stadtwald Böblingen	48.688682	9.052283	Trail
4	7	48.692019	9.107906	Waldrefugien	48.693604	9.108142	Forest

2.2 Data Preparing

Based on our business problem, the factors will influence our decision include:

- Locations of competitors.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
5	8	48.692019	9.117806	EDEKA Jäger	48.692259	9.122363	Supermarket
10	9	48.692019	9.127707	EDEKA Jäger	48.692259	9.122363	Supermarket
15	10	48.692019	9.137608	Euro Park	48.695377	9.139741	Shopping Mall
18	11	48.692019	9.147508	EDEKA Bauer	48.694187	9.143790	Supermarket
30	13	48.692019	9.167310	ALDI SÜD	48.694981	9.168974	Supermarket

- Locations of potential customers such as Asian restaurants and locations of public transports, which allows our supermarket to be easy to approach.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
14	10	48.692019	9.137608	Shell	48.694024	9.138979	Gas Station
31	13	48.692019	9.167310	Thang Long	48.688673	9.167308	Thai Restaurant
32	13	48.692019	9.167310	Kashmir	48.695733	9.170968	Indian Restaurant
41	14	48.692019	9.177211	Stuttgart Flughafen Bahnhof	48.691604	9.182186	Train Station
72	16	48.692019	9.197012	Stuttgart Airport Busterminal (SAB)	48.692009	9.197445	Bus Station
82	16	48.692019	9.197012	Fernbushaltestelle Stuttgart Flughafen / Messe	48.694136	9.192732	Bus Station
99	27	48.692019	9.305919	NORDSEE	48.693396	9.304835	Seafood Restaurant

Therefore, the data preparing step will include filtering needed data for those categories and cleaning in details irrelevant venues in these categories.

3. Methodology

The aim of this project is to detect the areas of Stuttgart which have high density of public transports and of restaurants, particularly Asian restaurants such as Japanese, Korean, Vietnamese, Thai and so on. In addition, we also put our effort on finding areas which have low density of supermarkets, shopping malls, food and drink shops. Our analysis limits the distance to the whole area of Stuttgart.

In first step we have collected the required data using Foursquare API, which involves:

- The coordinate of Stuttgart center and coordinated of all its neighborhoods.
- The data contains location and type (category) of every venue in each defined neighbourhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	0	48.692019	9.038601	Spatzensee	48.694011	9.043653	Playground
1	0	48.692019	9.038601	café im krankenhaus	48.690379	9.032925	Café
2	1	48.692019	9.048502	Initiativpark	48.692467	9.052749	Athletics & Sports
3	1	48.692019	9.048502	Stadtwald Böblingen	48.688682	9.052283	Trail
4	7	48.692019	9.107906	Waldrefugien	48.693604	9.108142	Forest

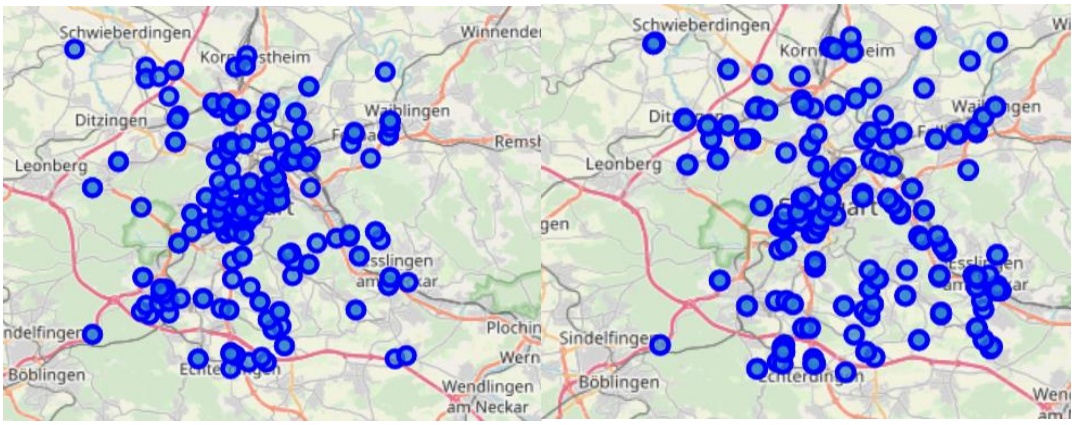
Second step in our analysis will be filtering of target locations of public transports, potential customers and potential competitors. After that, we calculate the distance from center of each neighbourhood to all favourite places and non-favourite places using haversine distance.

In third and final step we will focus on most promising areas that meet some basic requirements: we will take into consideration locations with no further than 0.5 kilometres to public transports or Asian restaurants, and we want locations without supermarkets, shopping mall and food and drink shops in radius of two kilometres. We will use reverse geocoding to request the address of those potential locations and visualize these locations again in the map.

4. Result

Our analysis shows that although there is a great number of restaurants and public transport in Stuttgart, namely 237 places, the number of competitors in this city are also very high, 238 competitors. Highest concentration of restaurants was detected around the city center. However, there are also a large number of restaurants and public transport in Stuttgart’s metropolitan area all in south, west, east and north of the city.

The distribution of competitors is also likely identical with the potential customer’s places, which is densest in the city urban area and spreading to the metropolitan areas with the lower density. However, in the west of the metropolitan area, there are very low density of competitors.



Favorite locations Non-favorite locations

After filtering out all favourite places and non-favourite places, we calculate the minimum distance from all neighbourhood centers to the competitors' locations as well as to the favourite locations. When the data is stored in a dataframe, it is very easy to further filtering step to set the requirement to the candidate locations. Here we consider only location within the radius of 0.5 kilometers to the potential customers and public transports station as well as having no competitors in the radius of 2 kilometers.

The addresses of these potential candidates were requested using reverse geocoding, therefore, there are no concrete house number but only the street name and area. These locations were plotted again to the map and showing the location in the west part of metropolitan area, where the distribution of competitors is very low.

	candidate_index	candidate_latitudes	candidate_longitudes	min_dist_to_competitor	min_dist_to_favourites	Address
121	121	48.716930	9.088104	2.207327	0.395566	Steinenbronn,Regierungsbezirk Stuttgart
411	411	48.779209	9.088104	2.719427	0.256864	Gerlingen,Regierungsbezirk Stuttgart
436	436	48.785437	9.048502	2.174846	0.409731	Gerlingen,Regierungsbezirk Stuttgart

Although the potential locations meet the requirement of potential customers and distribution of competitors, however, this result off course does not imply that those locations are optimal for opening an Asian supermarket. There may be some reasons that the number of both Asian restaurants and supermarkets are low here. The purpose of this project is only to provide an information of areas in Stuttgart which have Asian restaurant but having no nearby supermarkets.

5. Discussion

The weakness in the analysis is that we combine the location of public transports and potential customers into one category named favourite locations. This does not clearly show that the candidate locations are both closed to the potential customers and easy to approach. In the future, these two different favourite locations should be divided into two separate categories and finally using a condition to combine these two requirements together.

Furthermore, the distance to favourite places and competitors' locations in each neighbourhood are minimum distance. The average distance from the neighborhood's centers to these locations may reflect the situation more exactly, such as the minimum distance to one competitor in this neighbourhood is very low, however, it is only the competitor in the area, which is better than having within that neighbourhood many competitors in farther locations.

6. Conclusion

The purpose of this project is to find areas which have as many as possible Asian restaurant, bus or train stations but do not close to supermarkets or food shops. By analysing data from Foursquare, the needed features were selected. After the analysis, some potential areas were detected which meet the requirement of favourite and non-favourite places distribution. The project shaped the approach to the problem and further analysis can be conducted by varying the parameters to adjudge better locations. However, the result is not the optimal locations for a supermarket as it considers only one element which is location. To support better to the decision, other analysis should be conducted based on other important elements.