Find location for Japanese Restaurant in Beijing

1. Introduction

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening Japanese restaurant in Beijing, China. Since there are lots of restaurants in Beijing we will try to find locations that are not already crowded with restaurants. We're also particularly interested in areas with no Japanese restaurants in vicinity. We would also prefer locations as close to city center as possible, assuming that first 2 conditions are met.

We will use our data science method to generate a few most promising neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

2. Data

2.1 Data Source

Based on definition of our problem, factors that will influence our decision are:
- Number of existing restaurants in neighborhood
- Number of and distance to Japanese restaurants in the neighborhood, if any
- Distance of neighborhood from city center

We decided to use regularly spaced of locations, centered around city center, to define our neighborhood.

Following data sources will be needed to extract/generate the required information:
- Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those area obtained using Baidu Map API reserve geocoding
- Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API.
- Coordinate of Beijing center will be obtained using Baidu Maps API geocoding of well known Beijing location. (Tiananmen)

2.2 Data cleaning.

Firstly latitude and longitude coordinates will be created for centroids of our candidate neighborhoods, using Baidu Maps geocoding API.

Then a grid of area candidates will be created, which will be equally spaced, centered around city center and within ~6km from Tiananmen. Our neighborhoods will be defined as a circular

areas with a radius of 300 meters, so our neighborhood centers will be 600 meters apart.

To accurately calculate distances we need to create our grid of locations in Cartesian 2D coordinate system which allows us to calculate distances in meters. Then project those coordinates back to latitude/longitude degrees to be shown as Folium map. Baidu Map used BD09 spherical coordinate system, which need to be transferred to UTM Cartesian coordinate system.

The data can be visualized as blew folium map, including city center location and candidate neighborhood centers:
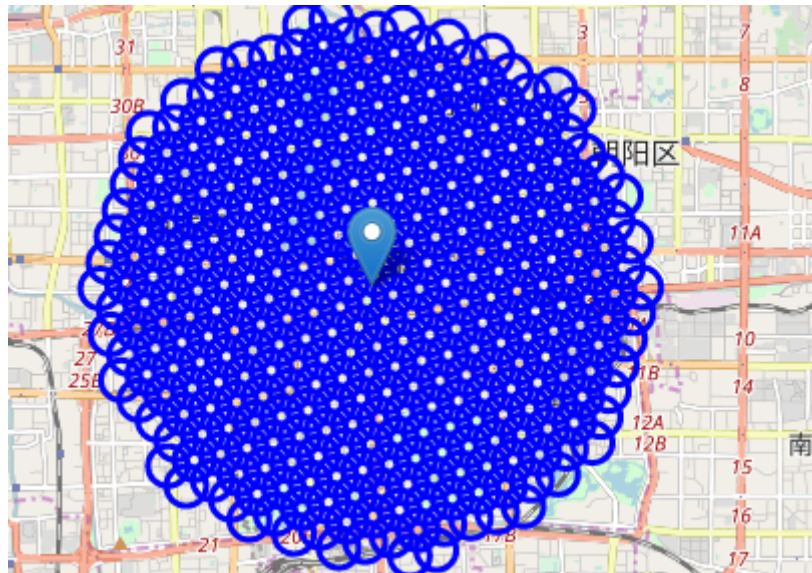

Figure1. Beijing Map with candidate neighborhoods

The Baidu Map API can be used to get approximate addresses of those locations, then put into a Pandas dataframe.

Foursquare API can be used to get information on restaurants in each location candidates. We're interested in venues in 'food' category, but only those that are proper restaurants – coffee shops, pizza places, bakeries etc. are not direct competitors so we don't care about those. So we will include in our list only venues that have 'restaurant' in category name, and make sure to detect and include all the subcategories of specific 'Japanese restaurant' category.

Then a map can be created to show all the collected restaurants in our area of interest, and Japanese restaurants in different color.
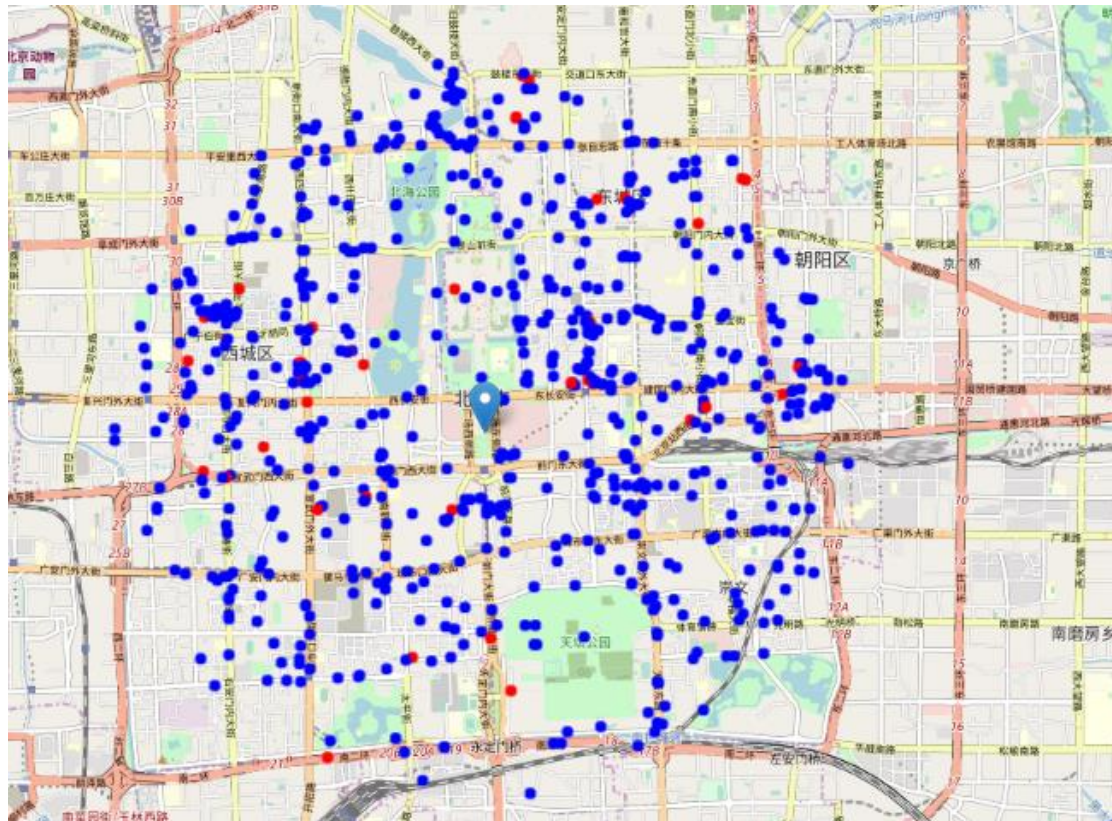
Figure2. Restaurant and Japanese Restaurant in Beijing Center

Finally, we are ready to use this data for analysis to produce the report on optimal locations for a new Japanese restaurant!

Methodology

In this project, we'll direct our efforts on detecting areas of Beijing that have low restaurant density, particularly those with low number of Japanese restaurants. We'll limit our analysis to area ~6km around city center.

In the first step we have collected the required data: location and category of every restaurant within 6km from Beijing Center. We have also identified Japanese restaurants.

Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of Beijing – we use heatmaps to identify a few promising areas close to center with low number of restaurants in general and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than 2 restaurants in radius of 250 meters, and we want locations without Japanese restaurants in radius of 400 meters. We will present map of all such locations but also create clusters using k-means clustering of those locations to identify general zones/neighborhoods/addresses which would be a starting point of final 'street level' exploration and search for optimal venue location by stakeholders.

Firstly we count the number of restaurants in every candidate, then calculate the distance to

nearest Japanese restaurant from every area candidate center. Then a Heatmap can be created to show the restaurants' density, borders of Beijing boroughs and a few circles indicating distance of 1km, 2km and 3km from Tiananmen are also shown in this map.
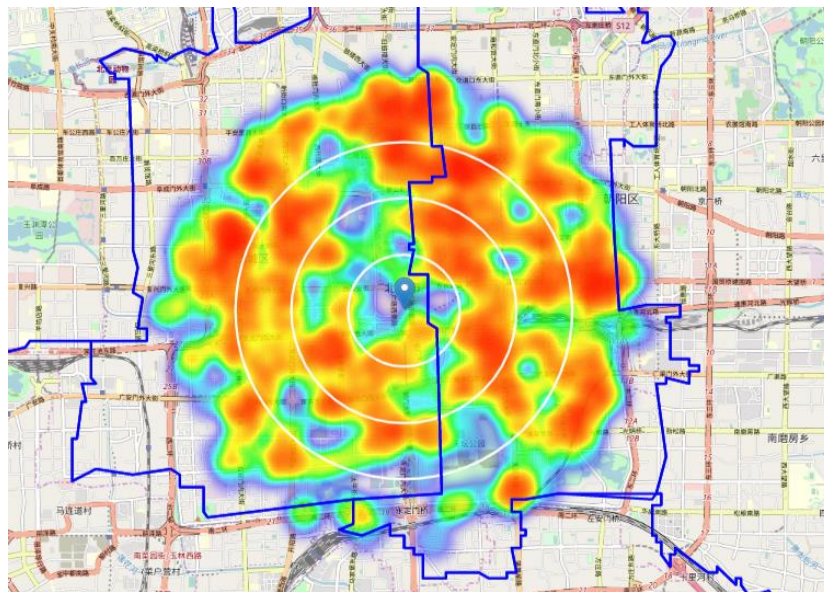

Figure3. Heatmap for Restaurant density

For better comparison, a heatmap specific for Japanese restaurant is also created.
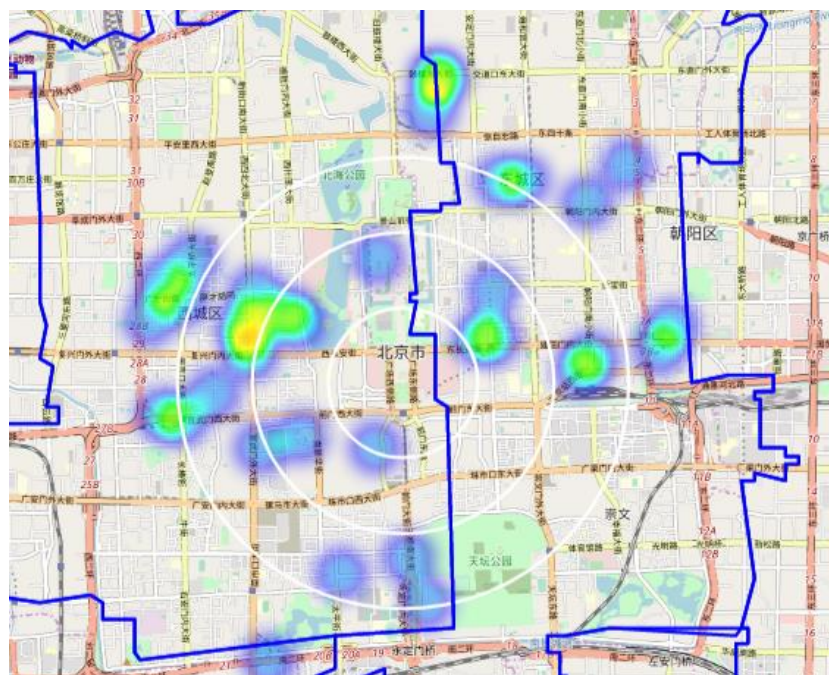

Figure4. Heatmap for Japanese Restaurant density

The map above is not so 'hot' but it also indicates higher density of existing Japanese restaurants directly west and east from Tiananmen, with closest pockets of low Japanese restaurant density positioned south-east and northwest.

Analysis from travelers guides, there is 1 famous park in northwest is Beihai Park. These are popular with tourists.

So then a new, more narrow region of interest is defined in the northwest area of Beijing

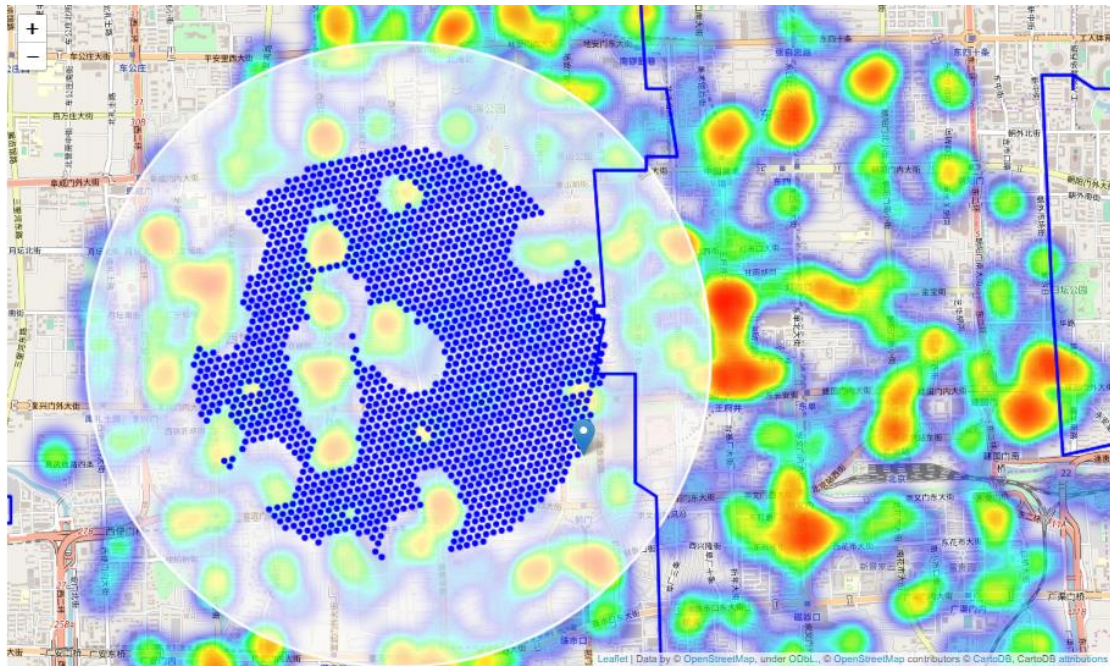center. And good locations can be showed in a form of heatmap below:

Figure5. Good location for new, narrow area

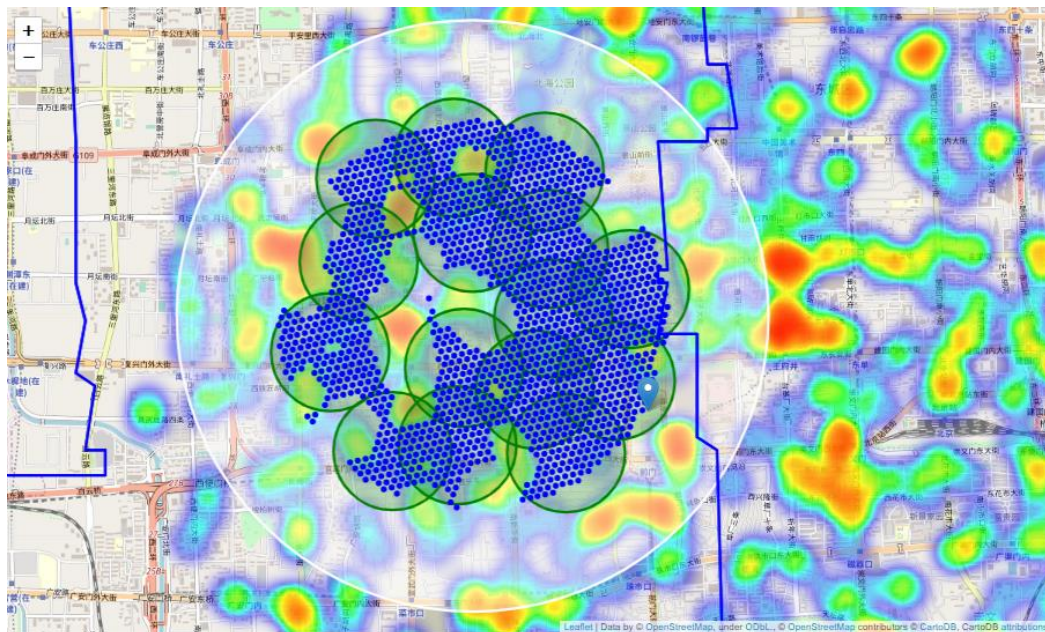K-mean will be used to cluster those locations to create centers of zones containing good locations.

Figure6. 15 Clusters for good locations.

And those address will be the final output of our analysis.

```
==========================================================
Addresses of centers of areas recommended for further analysis
==========================================================

北京市西城区萤光胡同9号                    -> 2.5km from 天安门
北京市西城区京畿道5号                      -> 3.2km from 天安门
北京市西城区西单北大街180-北门              -> 2.4km from 天安门
北京市西城区西便门东街6号                   -> 4.1km from 天安门
北京市西城区灵境胡同76号                    -> 3.3km from 天安门
北京市西城区金融大街21号                    -> 4.6km from 天安门
北京市西城区后铁厂胡同10号                  -> 0.6km from 天安门
北京市西城区胜利二巷16号                    -> 3.3km from 天安门
北京市西城区油坊胡同7号                     -> 1.7km from 天安门
北京市西城区复兴门桥                        -> 4.2km from 天安门
北京市西城区西章厂街54号                    -> 1.2km from 天安门
北京市西城区宣武门西大街99号                -> 2.4km from 天安门
北京市西城区北新华街13号                    -> 1.6km from 天安门
北京市西城区太平桥大街11号                  -> 4.1km from 天安门
北京市西城区上斜街12号                      -> 1.5km from 天安门
```

Result and Discussion

Our analysis shows that although there is a great number of restaurants in Beijing center, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected in northwest of Tiananmen, where is near the Beihai Park, which is also the famous place for tourist. So our attentions were focused on this area.

After directing our attention to this more narrow area of interest, we first create a dense grid of location candidates, those locations were then filtered so that those with more than 2 restaurants in radius of 250m and those with an Japanese restaurant closer than 400m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reversing geocoding.

Finally, result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance venues. The purpose of this analysis was only to provide information on area close to Beijing center but not crowded with existing restaurants, so the result does not imply those zones are actually optimal locations for new restaurants! Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion

Purpose of this project was to identify Beijing areas close to center with low number of restaurants(particularly Japanese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Japanese restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general areas that justify further analysis, and then generated extensive collection of locations which satisfy some basic

requirement regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest and addresses of those zone centers were created to be used as starting point for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommend zone, taking into consideration additional factors like attractiveness of each location, levels of noise to major roads, real estate availability, prices, social and economic dynamics of every neighborhoods.