

# Finding areas for an ethnic cuisine

Dmitry Ryabykh

07.03.2020

## 1. Introduction

If somebody is going to open a restaurant with ethnic cuisine, they should look for a location in areas, where many ethnic restaurants exist (which means that there is enough demand), but no restaurant of the cuisine they are going to open.

For example, I'm looking for a place for a new Thai restaurant. If I open it in area, where two other Thai restaurants exist, I risk meeting strong competition. If I open it on the city suburbs without other Thai restaurants nearby, my risk is low demand for such restaurant. However, if I open it in a neighborhood with several Chinese, Indian etc. restaurants, but no Thai restaurant, the chance of success will be higher.

Therefore, my application will create a map of a city and mark all areas with more than N different ethnic restaurants but no restaurant of selected ethnic cuisine.

I have chosen Thai restaurants in London as my example, but the application can be switched to any other big city and type of restaurant just by changing the link and the name of category.

## 2. Data acquisition and cleaning

I will use two sources of data:

**1. Areas.** For areas, I use list of areas from Wikipedia. Here is, for example, London:

[https://en.wikipedia.org/wiki/List\\_of\\_areas\\_of\\_London](https://en.wikipedia.org/wiki/List_of_areas_of_London)

It contains all areas with links to GeoHack service. I don't need the service. Instead I use the link itself to extract longitude and latitude from the URL. Unfortunately, some areas don't have the link – I exclude those areas from my list.

Eventually I create a list of areas with their latitude and longitude.

**2. Venues.** For the list of restaurants, I will use Foursquare.

I download the venues with their category. I use this category to check if the venues is an ethnic restaurant.

### **3. Methodology**

#### **Step 1 and 2. Initialize libraries**

First, I install Folium library to draw maps. Secondly, I import all other libraries necessary to work with data.

#### **Step 3. Load Wikipedia page with areas**

I download areas information from Wikipedia. In this example, I explore London. However, the same algorithm is applicable for many other big cities – they have the same wiki pages although some changes should be done in the page processing.

#### **Step 4. Load areas and their coordinates**

I use BeautifulSoup library to find the table with areas, load its name, then find the link to GeoHack and extract geographical coordinates from the link. The result is loaded into a pandas dataframe.

#### **Step 5. Count restaurants in each area**

Now I call Foursquare service to get all venues in each area. I do not store the list of venues. Instead I just count the venues with category equal to my category ('Thai Restaurant') and with category equal to one of other ethnic restaurants.

Thus, I have created a list of areas with two variables:

- Number of Thai restaurants
- Number of other ethnic restaurants

My idea is that a good area for a new Thai restaurant is an area, where ethnic restaurants are popular, so we find a lot of them, but no Thai restaurant is presented now.

In the application, I have defined the threshold for other ethnic restaurants as 10. During the development I tried other thresholds and chose the number that gives me about a dozen recommended areas.

#### **Step 6. Draw the map.**

Now we can put the data on the map. I draw the map of London and display all areas from my database marking recommended areas with red dots and all other areas with grey dots.

One can see recommended areas and explore them for possibility to open a new Thai restaurant.

## 4. Results

Here is the map that was produced for Thai restaurants in London for the minimum number of other ethnic restaurants set to 10:



We can see about a dozen areas, where ethnic restaurants are popular, but no Thai restaurant exists. Those are the places, where one should start exploring possibilities to open a new Thai restaurant.

## 5. Discussion

One of the parameters used in the analysis is the minimum number of ethnic restaurants to choose the area as having enough demand. I have set it to 10, as it gives me about a dozen recommended areas. If we get too many areas, it will be hard to explore them all. If we get too few, we don't have enough to choose from. For other cities or categories of restaurant this threshold could be different.

## 6. Conclusion

This application can help to find a balance between low demand for ethnic cuisine and high competition. This is an important step to find a good location for your restaurant. While some technical changes are necessary to apply it to other cities, the general principle will be the same for any city in the world.