Predicting Type and Location:
To open Restaurant
In Annemasse

Capstone Project - The Battle of the Neighborhoods (Week 2)

Wajeeha Farroqi

Introduction:

Background:

Opening a restaurant is one of the most daunting adventures in business, but it's also one of the most exhilarating and fulfilling. In the restaurant business today, there's greater competition and more financial challenges than ever before, whether they be food, labor, or real estate. Although restaurants across **France** are diverse, there are key similarities in the ones experiencing success. These restaurants are constantly innovating and adapting — leveraging technology, data, and new staff management practices to meet the changing expectations of their guests.

Our study will target one of the best business spots in France -- Haute-savoie -- province of France. Haute-Savoie is a region in the Alps of eastern France. Well Developed ,rich in beautiful sites, ski resorts--MONT BLANC--mountain range include popular Chamonix, with access to a large area of ski and snowboard runs. This province is rich with historical places, so it also has great value for tourists. In the view of geolocation, one of the communes of this province is close to the Switzerland border. This makes it more valuable places.These three properties--view,tourism and geolocation---make businessmen to look towards its commune for food business. All these properties will be clearly defined in one of the cities of Haute-savoie, ANNEMASSE.

Regardless of the approach, every restaurant defines success differently. This report outlines the key to open the type of restaurant in the best location according to the given parameters to help restaurateurs build their roadmaps for success in the coming age.

Problem-Statement:

According to three factors that affect the opening of the restaurant are site, tourism, and geolocation. Haute-savoie-Annemasse is one of the best sites for opening

the restaurant. The population of annemasse is 35,234, average income is 26,200 euros. The weather throughout the year is normal and seems too good for business point of view. The idea of this study is to help businessmen to open a new restaurant in Annemasse. In order to choose the right location by providing data about the income and population of each neighborhood as well as the competitors already present on the same regions.

Data Acquisition & Cleaning:

We analyse the data of population and income of Haute-Savoie. We decided to download Haute-savoie population data is available in this site *https://www.insee.fr/fr/statistiques/3689656#consulter* Each commune or neighbourhood revenue(income) of Haute-savoie are avalilable on this site *https://www.impots.gouv.fr/portail/statistiques*.

```
df.rename(columns={'Code commune':'Code'},inplace=True)
df.head(20)
```

	Code	Commune	Population	Income
0	74001	Abondance	1408	3797589
1	74002	Alby-sur-Chéran	2580	3070608
2	74003	Alex	1052	1363040
3	74004	Allèves	409	403747
4	74005	Allinges	4433	5816560
5	74006	Allonzier-la-Caille	2042	2461211
6	74007	Amancy	2579	4215269
7	74008	Ambilly	6302	11166567
8	74009	Andilly	870	1422424
9	74010	Annecy	126419	88034065
10	74012	Annemasse	35041	44661654
11	74013	Anthy-sur-Léman	2159	5533018
12	74014	Arâches-la-Frasse	1921	15539644
13	74015	Arbusigny	1102	1387194
14	74016	Archamps	2571	5829957
15	74018	Arenthon	1679	2400309
16	74019	Argonay	2866	4417814
17	74020	Armoy	1303	2324156
18	74021	Arthaz-Pont-Notre-Dame	1513	2865529
19	74024	Ayse	2132	1052201

Table1: list of communes by population,Income in haute-savoie

After sorting by population shown in Table2. And by income shown in Table3. We add another factor, weather. So we have best commune is annemasse .

```
df1=df.sort_values(by='Population',ascending=False)
```

```
df2= df1.head(5)
print(df2)
```

```
      Code           Commune  Population    Income
9    74010            Annecy      126419  88034065
246  74281  Thonon-les-Bains       35132  55026294
10   74012         Annemasse       35041  44661654
73   74081            Cluses       17371  22104479
224  74256        Sallanches       15902  25004859
```

Table2: list of top 5 communes by population in haute-savoie

```
df1=df.sort_values(by='Income',ascending=False)
```

```
df2= df1.head(5)
print(df2)
```

```
      Code              Commune  Population    Income
9    74010               Annecy      126419  88034065
246  74281     Thonon-les-Bains       35132  55026294
155  74173               Megève        3123  46830727
48   74056  Chamonix-Mont-Blanc        8759  45686039
10   74012            Annemasse       35041  44661654
```

Table3: list of top 5 communes by Income in haute-savoie

So , in fig1. Shows annemasse has average population and income and the factor which differentiate it from the rest of commune is the weather. Throughout the year, the weather is pleasant here.

The data is preprocessed by taking the population column and code of all communes in Haute-savoie 2016. The income column is taken from impot.fr website . I merge these columns in one table. The data analysis verifies that annemasse is our target city.

Following data sources will be needed to extract/generate the required information:

number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API

coordinates of commune are extracted from geocode API

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 areas will be obtained using Google Maps API reverse geocoding
- number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API

```
df2.plot('Commune',['Population','Income'],kind='line')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x24dab5439d0>
```

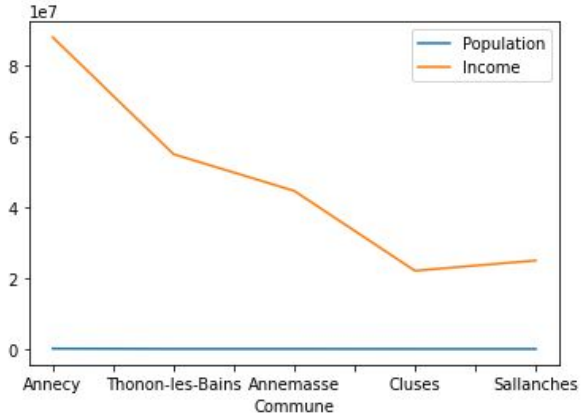

Fig1: graph shows top 5 communes by population in haute-savoie

Methodology:

In this project we will direct our efforts on detecting areas of Annemasse that have low restaurant density, particularly those with low number of sushi,rest type of restaurants. We will limit our analysis to areas ~6km around the city center.
In the first step we have collected the required data: location and type (category) of every restaurant within 6km from Annemasse center (Marie d'annemasse). We have also identified different types of restaurants (according to Foursquare categorization).
Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of annemasse - we will use heatmaps to identify a few promising areas close to the center with a low number of different types of restaurants in general.
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 two restaurants in radius of 250 meters, and 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 should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders. Before we cluster the 364 location, we need to explore the food category. We map the information on the map, so that we can explore the food information clearly.

Analysis:

Let's perform some basic exploratory data analysis and derive some additional info from our raw data. First let's count the number of restaurants in every area candidate in fig 2.The fig2. shows the type of restaurants already in Annemasse. The number of restaurants of French, Italian,Indian, chinese are high. Sea food, fast food restaurants are very low in number. In fig 3. Shows the one of high peak restaurants in annemasse which is near the border and away from the center.

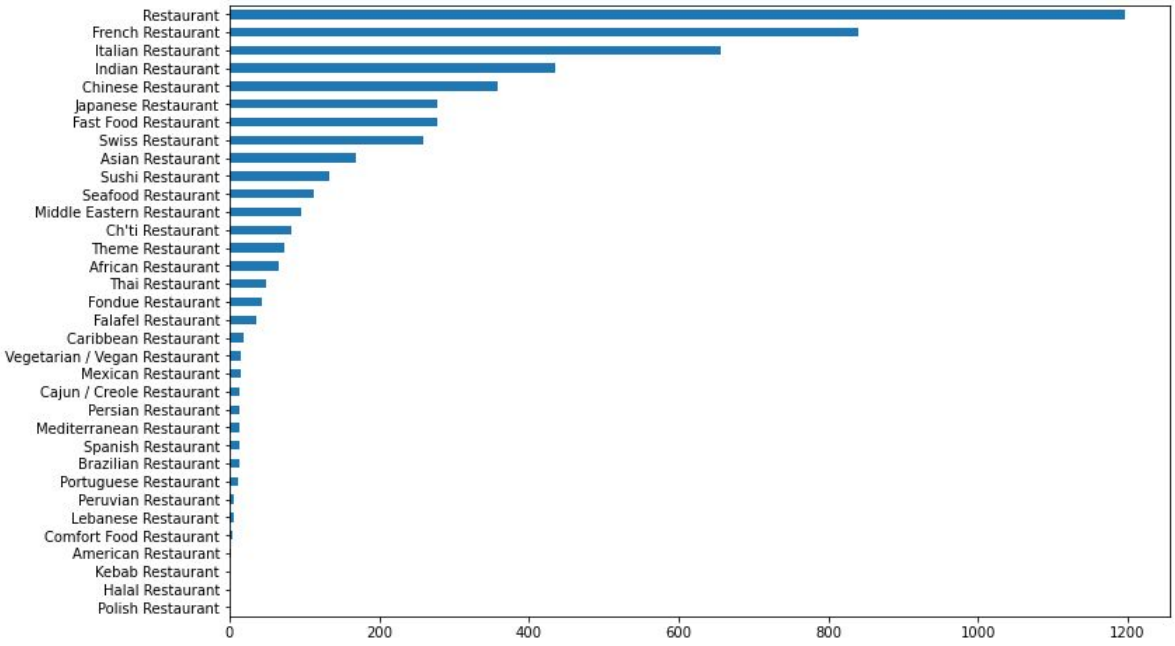

Fig2: Types of Restaurants in Annemasse

Let us now cluster those locations to create centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis.

Results and Discussion:

Our analysis shows that although there is a great number of restaurants in Annemasse City (~2000 in our initial area of interest which was 12x12km around Annemasse), the French restaurant is the main type close to the city center. So, if we want to choose an appropriate location to open a new restaurant, we choose from the target.

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

Chemin des Sillons 7, 1226 Thônex, Switzerland     => 2.5km from Marie
75 Route de Thonon, 74100 Annemasse                => 2.6km from Marie
47 Rue de la Bernadette, 74560 Monnetier-Mornex    => 4.9km from Marie
Route de La-Louvière 71, 1243 Presinge, Switzerland => 4.5km from Marie
4 Route de Collonges, 74100 Vétraz-Monthoux        => 2.0km from Marie
Chemin de Sionnet 55, 1254 Jussy, Switzerland      => 4.7km from Marie
Route de Sous-Moulin 65, 1226 Thônex, Switzerland  => 4.8km from Marie
Route de Choulex 220, 1244 Choulex, Switzerland    => 4.7km from Marie
265 Chemin de Nanteux, 74930 Reignier-Esery        => 4.8km from Marie
Chemin de la Blonde 3, 1253 Vandœuvres, Switzerland => 5.0km from Marie
Route de Presinge 58, 1241 Puplinge, Switzerland   => 1.7km from Marie
17 Chemin de la Prairie, 74100 Vétraz-Monthoux     => 4.8km from Marie
205 Chemin du Pont du Loup, 74560 Monnetier-Mornex => 4.5km from Marie
2 Rue de Montréal, 74100 Ville-la-Grand            => 5.0km from Marie
21 Rue René Cassin, 74240 Gaillard                 => 2.3km from Marie
```

Fig3: List of largest number of potential new restaurant

Here are pockets of low restaurant density fairly close to the city center. Highest concentration of restaurants was detected near the switzerland border , so we focused our attention to the center and south-east of the center of the annemasse within 5 km. but our attention was focused on the center of the city which offers a combination of popularity among tourists, closeness to the train station, strong socio-economic dynamics and a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest.Those location candidates were then clustered to create zones of interest which contain the greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones shown in the fig 3. containing the largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Berlin center but not crowded with existing restaurants - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. 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 Annemasse areas close to the center with a low number of restaurants (particularly fast food) in order to aid stakeholders in narrowing down the search for the optimal location for a new restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points 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 recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.