Predicting the best place for a restaurant opening in Lausanne (Switzerland)

Nicolas Vuille-dit-Bille

22 March 2020

1. Introduction

1.1 Background

To open a new restaurant in a new city, it's difficult to choose the right place to start the business. There are a lot of parameters to take into account in order to succeed to run a new business. The food, service but also the location has to be chosen in a meaningful way in order to establish a productive business. The location of the restaurant should take into account different parameters like parking availability, visibility and neighbourhood competition. With Foursquare location data, it will be possible to recommend a specific location adapted to a new restaurant opening.

1.2 Problem

The right location for a specific restaurant is essential in order to make benefits rapidly and avoid competition problems. The environment around the restaurant location should also give information about the customers type and the consumption trend in the neighbourhood.

1.3 Interest

The utilization and the analysis of geospatial data concerning neighbourhood and its most frequented places could be interesting for example for some customers which would like to invest in a new restaurant and therefore they would like to find the best appropriate location to start their business.

2. Data acquisition and cleaning

2.1 Data sources

To recommend a location for an opening of a new restaurant, the data from data.geo.admin.ch have been used in order to have geodata of the city of the interest like the commune, postcode, longitude and latitude for each neighbourhood (locality).

2.2 Data cleaning

Data has been scrapped from data.geo.admin.ch as a zip file (https://data.geo.admin.ch/ch.swisstopo-vd.ortschaftenverzeichnis_plz/PLZO_CSV_WGS84.zip). This zip file has been unzipped and inserted into a dataframe with six features : Ortschaftsname, PLZ, Zusatzziffer, Gemeindename, BFS-Nr, Kantonskürzel, E, N, Sprache. These features have been renamed respectively in Locality, Postcode, Amendment, Commune, BFS-Nr, Canton, Latitude, Longitude, Language.

2.3 Feature selection

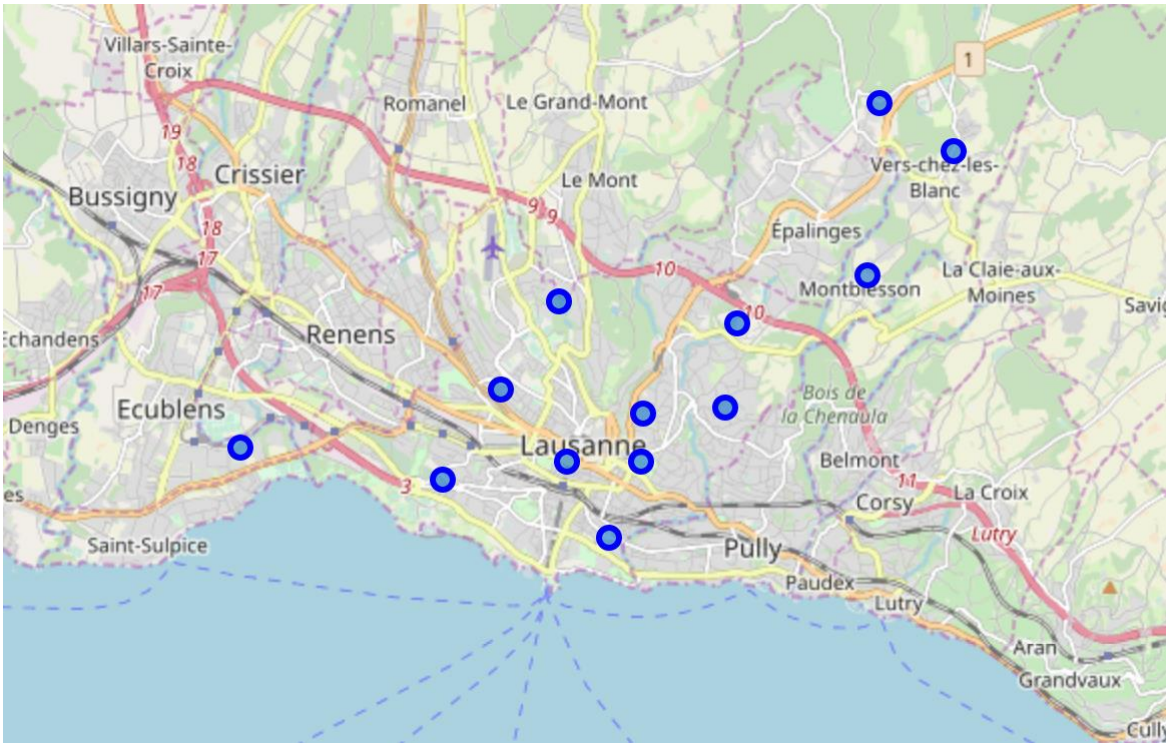
The analysis was focused on the city of Lausanne. So, only information related to this city was extracted and inserted into a new dataframe. Some features (Amendment, BFS-Nr, Canton and Language) were not useful for the analysis and were therefore removed from the dataframe.

Table 1. Feature and Lausanne data selection during data cleaning.

	Locality	Postcode	Longitude	Latitude
0	Lausanne 25	1000	6.683444	46.562237
1	Lausanne 26	1000	6.696216	46.556483
2	Lausanne 27	1000	6.681465	46.541743
3	Lausanne	1003	6.630034	46.520004
4	Lausanne	1004	6.618678	46.528480
5	Lausanne	1005	6.642500	46.519859
6	Lausanne	1006	6.637110	46.510849
7	Lausanne	1007	6.608606	46.517754
8	Lausanne	1010	6.658920	46.536143
9	Lausanne	1011	6.642880	46.525635
10	Lausanne	1012	6.656931	46.526342
11	Lausanne	1015	6.574090	46.521485
12	Lausanne	1018	6.628646	46.538725

3. Exploratory Data Analysis

3.1 Data representation on a map



3.2 Cluster analysis of different Lausanne areas


```
# Cluster 1
lausanne_merged.loc[lausanne_merged['Cluster Labels'] == 0, lausanne_merged.columns[[1] + list(range(5, lausanne_merged.shape[1]))]]
```

Out[37]:

	Postcode	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	1004	Bus Stop	Grocery Store	Supermarket	Bus Station	Jewish Restaurant	Restaurant	Italian Restaurant	Sporting Goods Shop	Stadium	Steakhouse
7	1007	Bus Station	Snack Place	Restaurant	Supermarket	Hotel	Grocery Store	Gas Station	Construction & Landscaping	Park	Gym
8	1010	Gym	Gas Station	Bus Station	Sushi Restaurant	Wine Bar	Ethiopian Restaurant	Cupcake Shop	Deli / Bodega	Department Store	Dessert Shop
12	1018	Grocery Store	Bus Stop	Bus Station	Supermarket	Wine Bar	Ethiopian Restaurant	Cupcake Shop	Deli / Bodega	Department Store	Dessert Shop

```
# Cluster 2
lausanne_merged.loc[lausanne_merged['Cluster Labels'] == 1, lausanne_merged.columns[[1] + list(range(5, lausanne_merged.shape[1]))]]
```

:

	Postcode	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
11	1015	Stadium	Deli / Bodega	Light Rail Station	College Cafeteria	Wine Bar	Creperie	Cupcake Shop	Department Store	Dessert Shop	Diner

```
# Cluster 3
lausanne_merged.loc[lausanne_merged['Cluster Labels'] == 2, lausanne_merged.columns[[1] + list(range(5, lausanne_merged.shape[1]))]]
```

:

	Postcode	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	1003	Bar	Café	Italian Restaurant	French Restaurant	Plaza	Japanese Restaurant	Swiss Restaurant	Lounge	Burger Joint	Gym
5	1005	Bar	Pizza Place	Nightclub	Hotel	Italian Restaurant	French Restaurant	Park	Bakery	Supermarket	Burger Joint
6	1006	Café	Park	Skating Rink	Museum	Pharmacy	Art Museum	Bakery	Bar	Plaza	Pool
9	1011	Thai Restaurant	Hotel	Sushi Restaurant	Vietnamese Restaurant	French Restaurant	Italian Restaurant	Massage Studio	Metro Station	Museum	Pizza Place

```
# Cluster 4
lausanne_merged.loc[lausanne_merged['Cluster Labels'] == 3, lausanne_merged.columns[[1] + list(range(5, lausanne_merged.shape[1]))]]
```

:

	Postcode	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	1000	French Restaurant	Swiss Restaurant	Spa	Restaurant	Gas Station	Food Court	Wine Bar	Creperie	Cupcake Shop	Deli / Bodega
1	1000	French Restaurant	Swiss Restaurant	Spa	Restaurant	Gas Station	Food Court	Wine Bar	Creperie	Cupcake Shop	Deli / Bodega
2	1000	French Restaurant	Swiss Restaurant	Spa	Restaurant	Gas Station	Food Court	Wine Bar	Creperie	Cupcake Shop	Deli / Bodega

```
# Cluster 5
lausanne_merged.loc[lausanne_merged['Cluster Labels'] == 4, lausanne_merged.columns[[1] + list(range(5, lausanne_merged.shape[1]))]]
```

:

	Postcode	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
10	1012	Department Store	Bakery	Sushi Restaurant	Supermarket	Fast Food Restaurant	Wine Bar	Ethiopian Restaurant	Creperie	Cupcake Shop	Deli / Bodega

4. Conclusions

According the restaurant competition, cluster 3 and cluster 4 should be avoided to start a new restaurant. Cluster 1, 2 and 5 are better place to start a new restaurant business because there are not so much famous restaurants in these areas.