

Capstone Project: Thai Restaurant in Switzerland

June, 2020

1 Introduction

The business problem presented here is about getting a new Thai restaurant opened in Switzerland. When living in Switzerland one knows that there are already many Thai restaurants in the country, because Swiss people enjoy eating in restaurants (no matter whether very simple or expensive) and the spices and flavors of the Thai cuisine seem to be a favorite in the country.

Someone wanting to open a Thai restaurant would wonder what are the places in the country that would make sense from a business perspective considering:

- restaurant should be in a densely populated region with strong economic activity and relatively high income
- restaurant should not be too close to already existing Thai restaurants

It would be interesting to know whether there are still some well populated areas in Switzerland still with no Thai restaurant around.

2 Data

The data (and their sources) that should help answer the questions posed in the introduction section:

1) GDP per capita, from the Swiss Federal Office of Statistics (by cantons):

<https://www.bfs.admin.ch/bfs/de/home/statistiken/volkswirtschaft/volkswirtschaftliche-gesamtrechnung/bruttoinlandprodukt-kanton.assetdetail.10647589.html>

2) Income per capita, from the Swiss Federal Office of Statistics (by cantons):

<https://www.bfs.admin.ch/bfs/de/home/statistiken/wirtschaftliche-soziale-situation-bevoelkerung/einkommen-verbrauch-vermoegen/haushaltsbudget.assetdetail.10867286.html>

3) Data about population in the various municipalities in Switzerland, with number of inhabitants per municipality. This information can be obtained from this site: <https://opendata.swiss/de/group/population>

Additionally, the following will be used with Python in order to support the project:

4) *Nominatim Geopy* module to locate the coordinates of municipalities.

5) *Foursquare* data to identify already existing Thai restaurants in Switzerland and their respective location.

3 Methodology

In addition to the data and libraries already mentioned above, this project makes use of the following libraries to understand, prepare, model and visualize data:

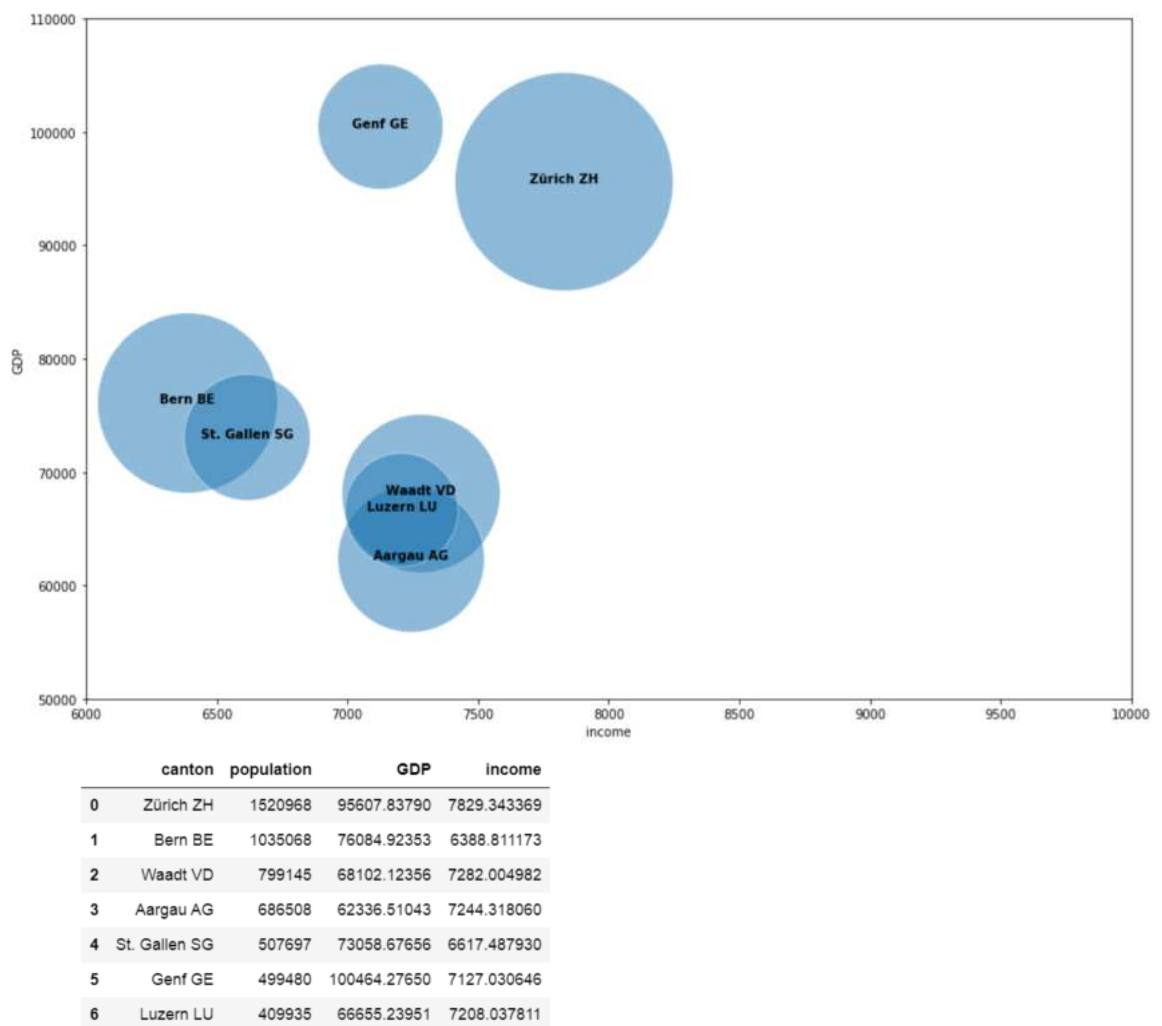
- Numpy
- Pandas
- Json
- Matplotlib
- Seaborn
- Folium

Working with the data mentioned in section 2, the following was required for data preparation:

- Bring raw data into pandas dataframe
- Describe data for initial understanding
- Cleansing data and look for missing data (pandas "isnull()")
- Rearrange and concatenate data providing the main dataframes that will be discussed further on

Finally some decisions had to be made based on information collected:

- Scope: in order to narrow the search to a meaningful, but still promising, amount of municipalities, a decision was made that it would be enough to look only for places within the 7 most populous cantons of Switzerland, which all have significant populations and also income/GDP (per person, shown in this analysis in Swiss Francs) that certainly could justify the investment
- The spread/bubble chart show these 7 cantons (X=income p.p., Swiss Francs, Y=GDP p.p., Swiss France, Bubble Size = proportional to population size with details on table further down)



The next decision was about selecting only municipalities (within the 7 cantons mentioned above) that already have a population greater than 5'000 people as a threshold. This reduced further the scope of our search to 251 potential municipalities (below is a print of the dataframe shape)

(251, 7)

Finally, it was time to obtain the location of all Thai restaurants (venues, using Foursquare) within 5 km of any of the 251 municipalities pre-selected.

Then it was possible to classify the municipalities with a Thai restaurant nearby according to the number of restaurants within less than 5 km:

Classification of Municipality	Numbers of restaurants within 5 km
A	1 – 3
B	4 – 7
C	8 – 10
D	11 – 25
E	26+

Categories (5, object): [A: 1-3 < B: 4-7 < C: 8-10 < D: 11-25 < E: 26+]

This classification showed that, for example there were 105 municipalities with 1 to 3 restaurants nearby:

```
A: 1-3      105
B: 4-7      69
C: 8-10     16
D: 11-25    12
E: 26+      8
Name: Classification, dtype: int64
```

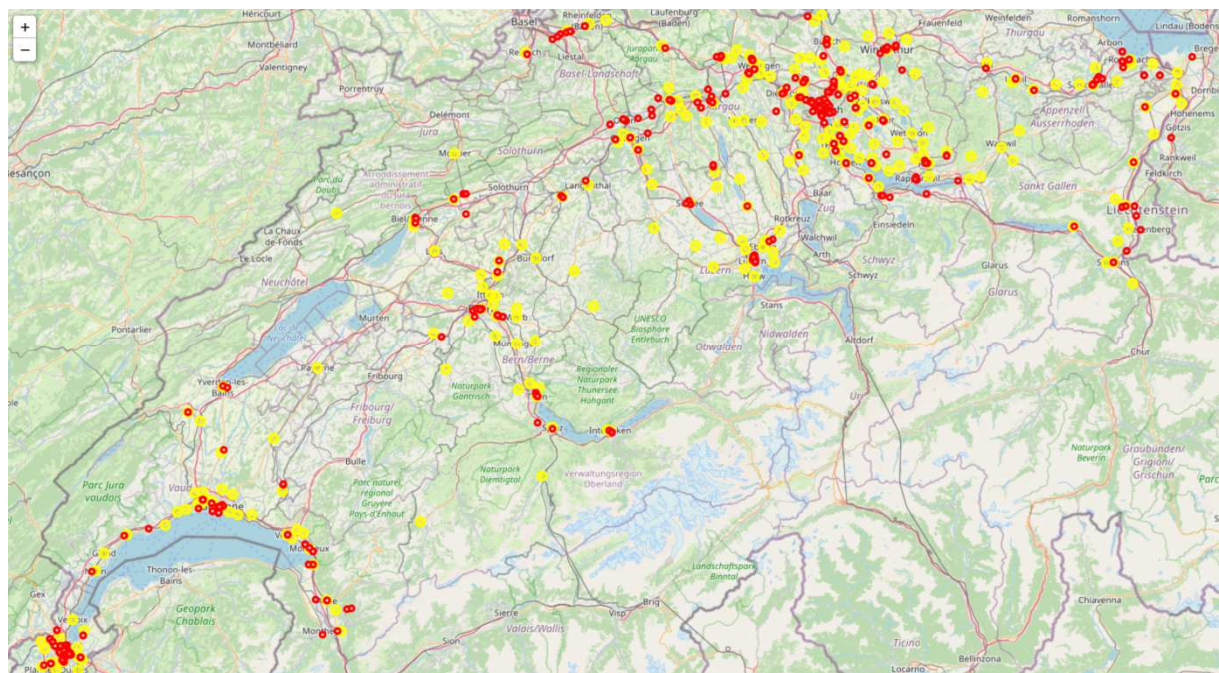
By inference it also showed that 41 municipalities (out of 251 pre-selected) had **no** Thai restaurant within reach of 5 km.

In the final step, the pre-selected municipalities and venues (Thai restaurants) were plotted on the map of Switzerland to support decision making.

4 Results

On the map:

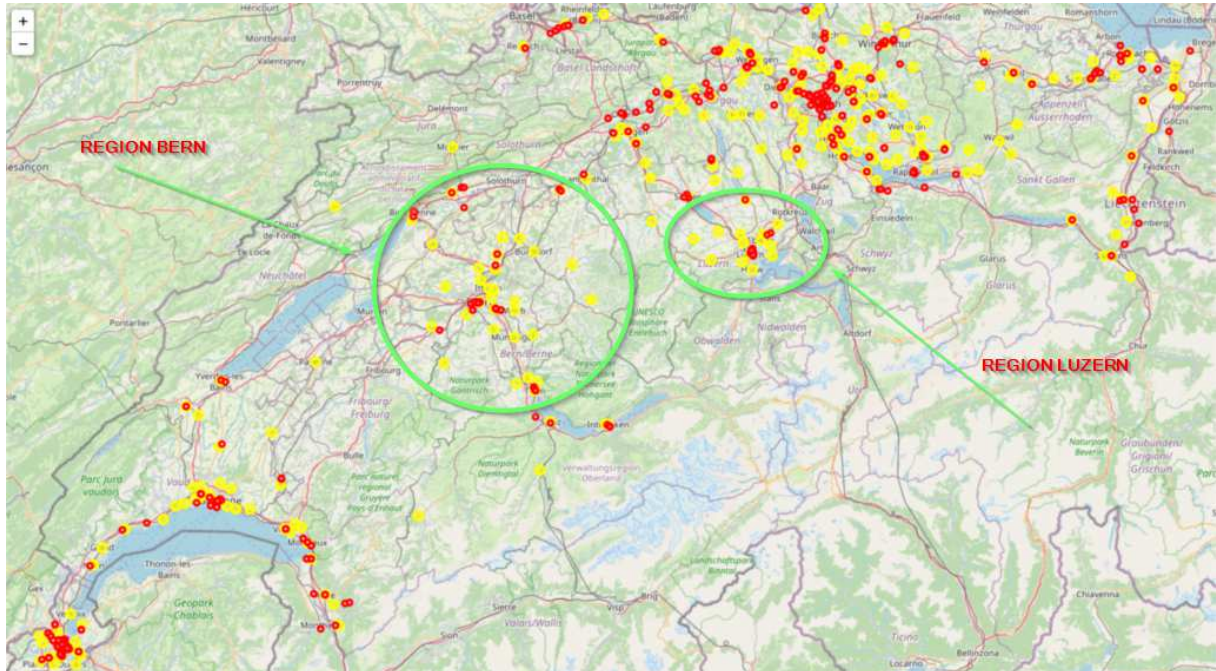
- yellow dots represent the 251 pre-selected municipalities
- Red dots represent all Thai restaurants within 5 km of municipalities



5 Discussion

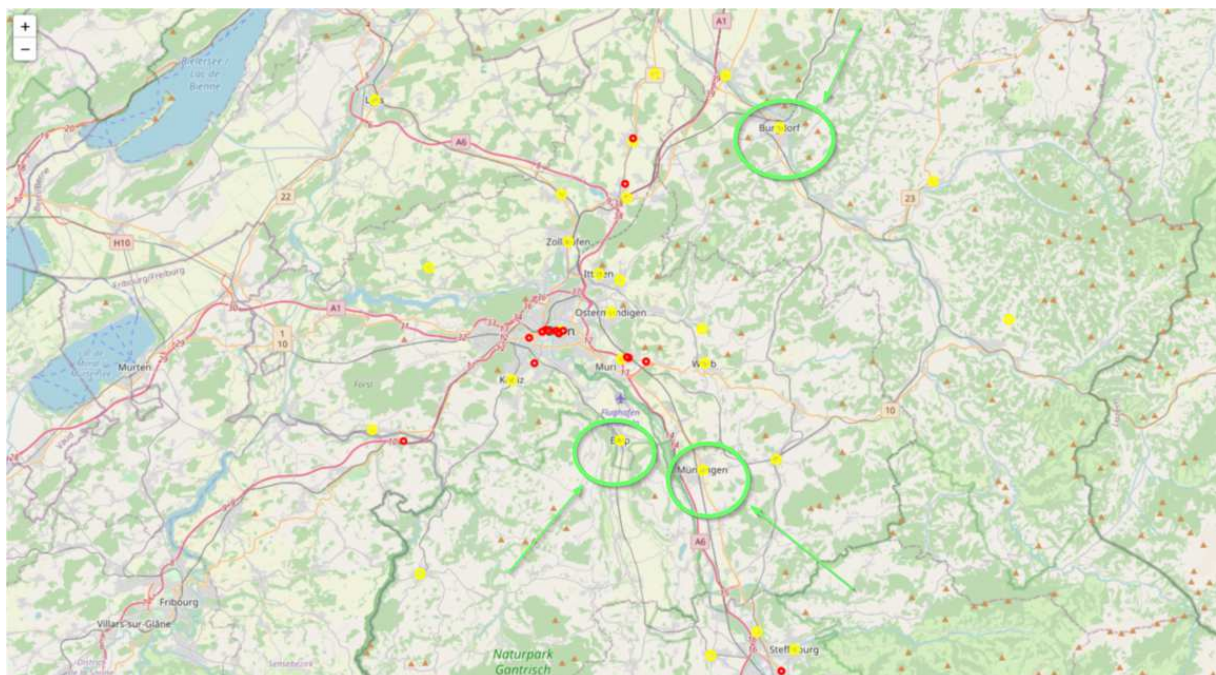
It is clear that there are some regions, such as Zurich and Geneva with very high concentration of Thai restaurants.

However, it is also clearly visible that some areas, e.g. around Bern or Luzern, where there is a concentration of relatively large municipalities with not that many restaurants nearby.

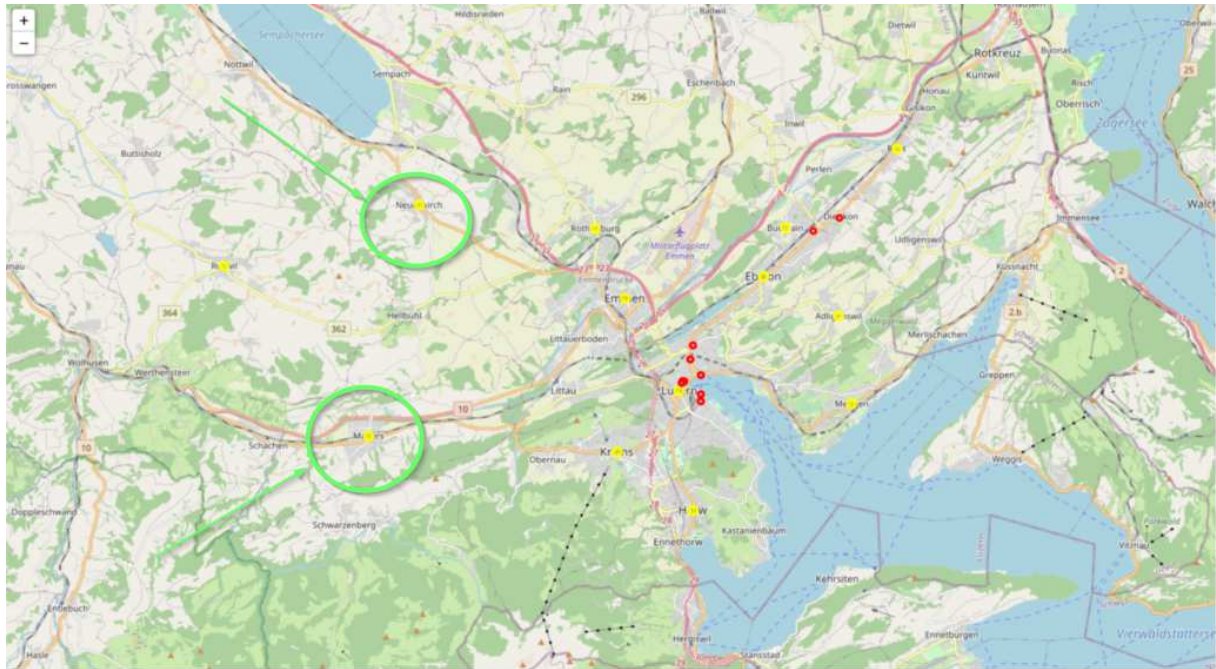


By examining these two regions a bit closer, it is possible to identify good candidates for a new Thai restaurant.

In the region of Bern, these candidates would be e.g. the municipalities of Belp, Münsingen and/or Burgdorf:



In the region of Luzern, these candidates would be e.g. the municipalities of Malters and/or Neunkirch:



6 Conclusion

The use of data analysis to identify possibilities for a new business location has proved to be possible, thanks to the amount of data available in the internet and tools (IBM Watson Studio, libraries, API's such as Foursquare, etc). Nevertheless it still requires a lot of time to prepare and model data. Additionally it is still fundamental to acquire understanding of the business and the problem one is trying to solve in order to achieve good results.