

Switzerland's secrets

Michael Schuemperli

07-May-2021

1. Business Problem

Switzerland is a small country in Central Europe, famous for watches, chocolate and as tax haven. Switzerland also features breathtaking landscapes – a great portion of the country lies in the alps – and tourism is an important economy.

As a tourist, you may want to get an idea which regions would be interesting to visit and what kind of venues you would come across there. Where to go for skiing and hiking, are there hotels, restaurants, other types of recreation? Where will you be in the crowds and where can you expect untouched nature?

As someone who plans to go live and work in Switzerland, you may be interested in getting an impression of where the Swiss live and work, where GDP is higher or lower and what venues to expect there.

2. Data

Four different datasets are combined in this analysis:

- Swiss location data
- Population density
- GDP
- Foursquare location data

Firstly, Switzerland will be divided into regions. Luckily, the Swiss did the job already, they subdivided their country into 26 regions (cantons).

Swiss location data will be downloaded from GADM.org, level 1 providing the canton border data. This data is going to be the basis for the maps to be drawn. GADM location data can be downloaded as GeoJSON file and can be directly used with folium to draw maps.

In this work, Swiss location data will be combined with Population density and GDP data for each canton into two different choropleth maps of the country.

GDP and population density data will be downloaded from the Swiss Federal Statistical Office in xlsx format. The dataset will be cleaned and readied for drawing the needed maps. That not only means

to remove empty cells, columns and lines from the original dataset, but also to add the appropriate key for drawing the choropleth maps.

Secondly, Foursquare location data will be used to characterize and group the 26 cantons. The grouping will again be combined with each of the two choropleth maps. Every group of cantons will be characterized by the top 10 venues found there.

Switzerland's area covers 41285 km². As stated above, the country is divided into 26 cantons, giving an average area of 1588 km² per canton. Assuming a circle, the radius was 22.5 km per canton. In real, the cantons have different sizes, and there are small and very large cantons. Therefore, the radius used in the Foursquare API will be 40 km, accepting some overlap in some areas in favor of creating too many gaps.

3. Methods

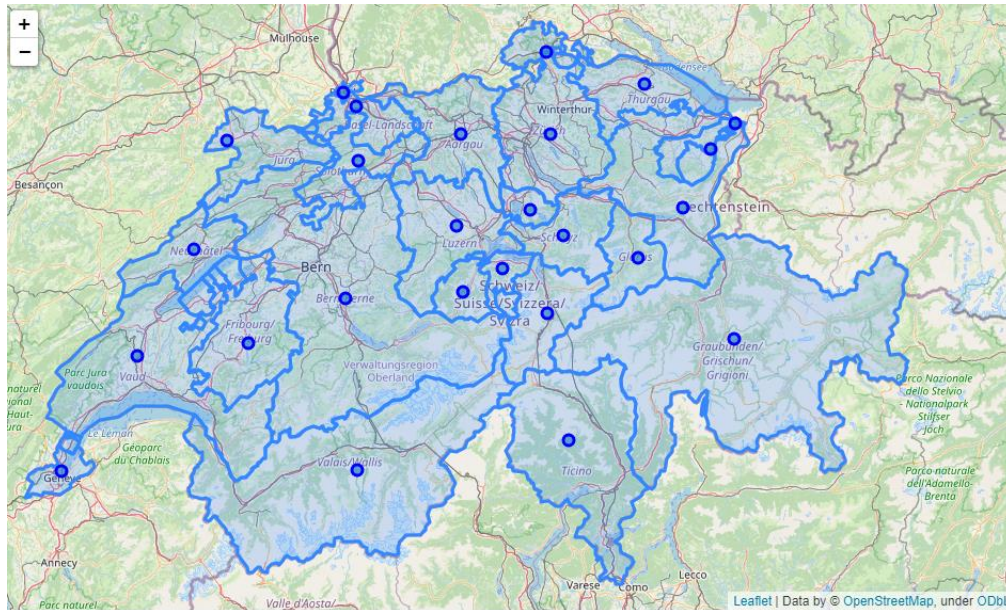
A dataframe needs to be prepared, containing all the relevant data:

- Canton names
- GDP
- Population Density
- Latitude
- Longitude

The xlsx from the Swiss Federal Statistical Office contains more data than needed. The xlsx is loaded into Jupyter as is and then sliced and cleaned.

Geographical data (latitude and longitude) for the cantons is added to the dataframe via Nominatim. The input is the column with the canton names. However, in some cases, the canton names are not unique, since some cantons' main cities are called the same as the canton, and in two instances, the canton names point to geographical areas outside of Switzerland. Before Nominatim delivers the geographical data of the canton centers, the canton names have to be adapted.

In order to check whether the geographical data of the canton returned the canton centers, the locations are mapped, indicated by the blue dots:

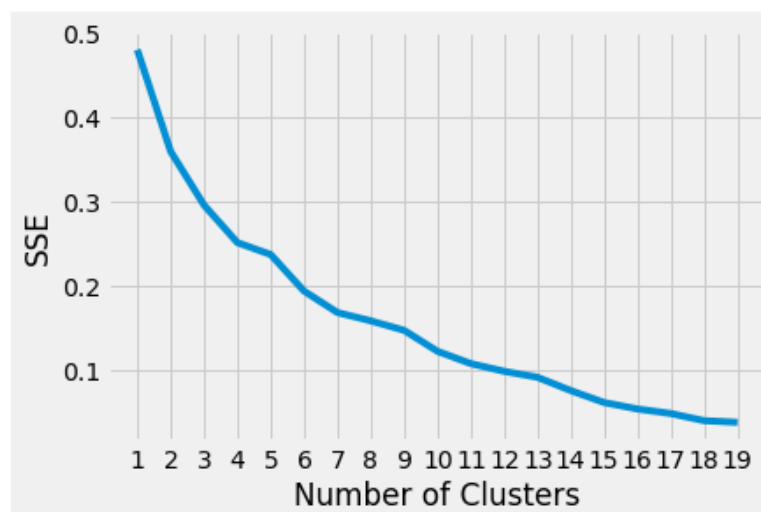


The aim is to draw maps of Switzerland giving an impression on GDP, population density and canton clustering according to similarities. As pointed out above, GeoJSON data is loaded containing canton border data. The key used for drawing the Choropleth maps is HASC_1, a unique code for every canton. In order to combine the geographical data with the dataframe, the HASC_1 key has to be added as column 'code'.

Foursquare location data is parsed for every canton with a radius of 40 km around the geographical center. 2600 venues in 189 unique categories are returned. These venues are sorted by category and the resulting dataframe is sliced to the Top 10 venues in each canton.

Then, the cantons are clustered with Kmeans. The first step for clustering is determination of the number of clusters. This is achieved with the Elbow method:

Kmeans is iterated and the sum of the squared error (SSE) is calculated. Kmeans is limited to 10 runs of 300 iterations each, which should be by far enough to allow the centroids to converge. SSE is then plotted against number of clusters. The optimum number of clusters can be read from that graph as the elbow:



In this case, the elbow cannot be easily estimated by eye. Kneelocator is used to calculate the elbow from the data:

```
#calculate number of clusters with KneLocator
kl = KneLocator(
  range(1, 20), sse, curve="convex", direction="decreasing"
)

kclusters = kl.elbow
kclusters
```

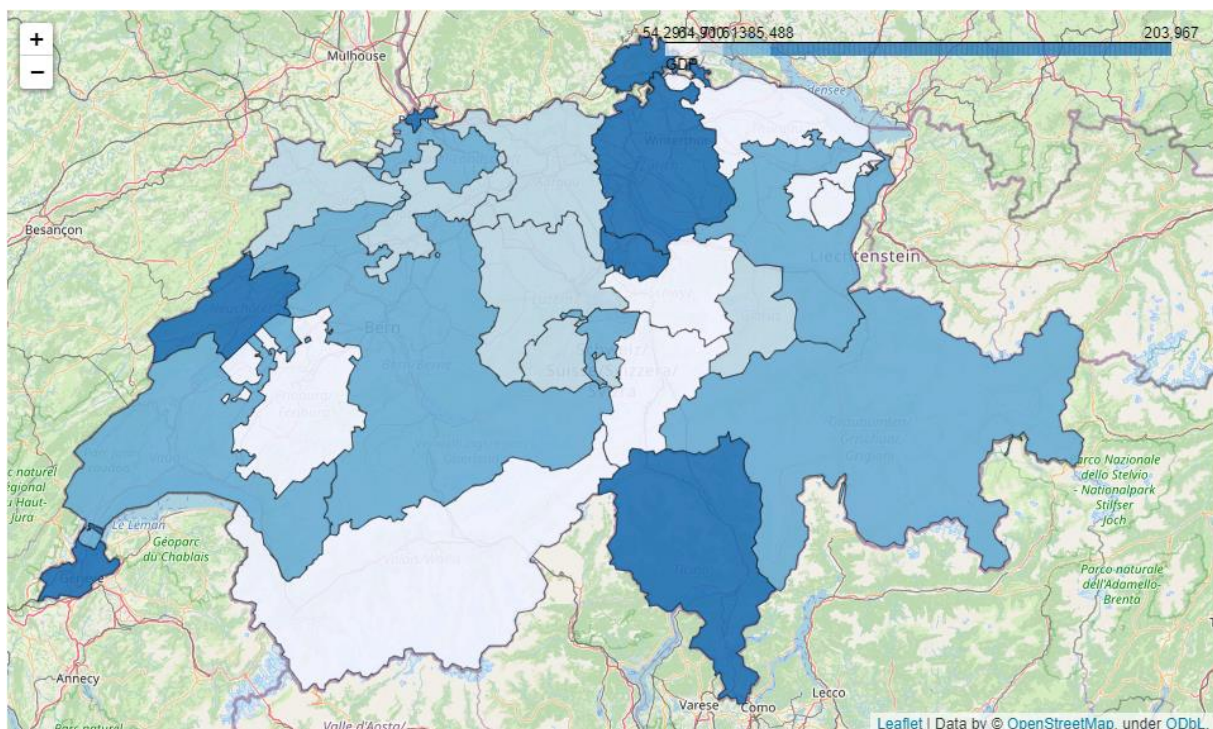
7

The number of 7 clusters is used for the Kmeans clustering.

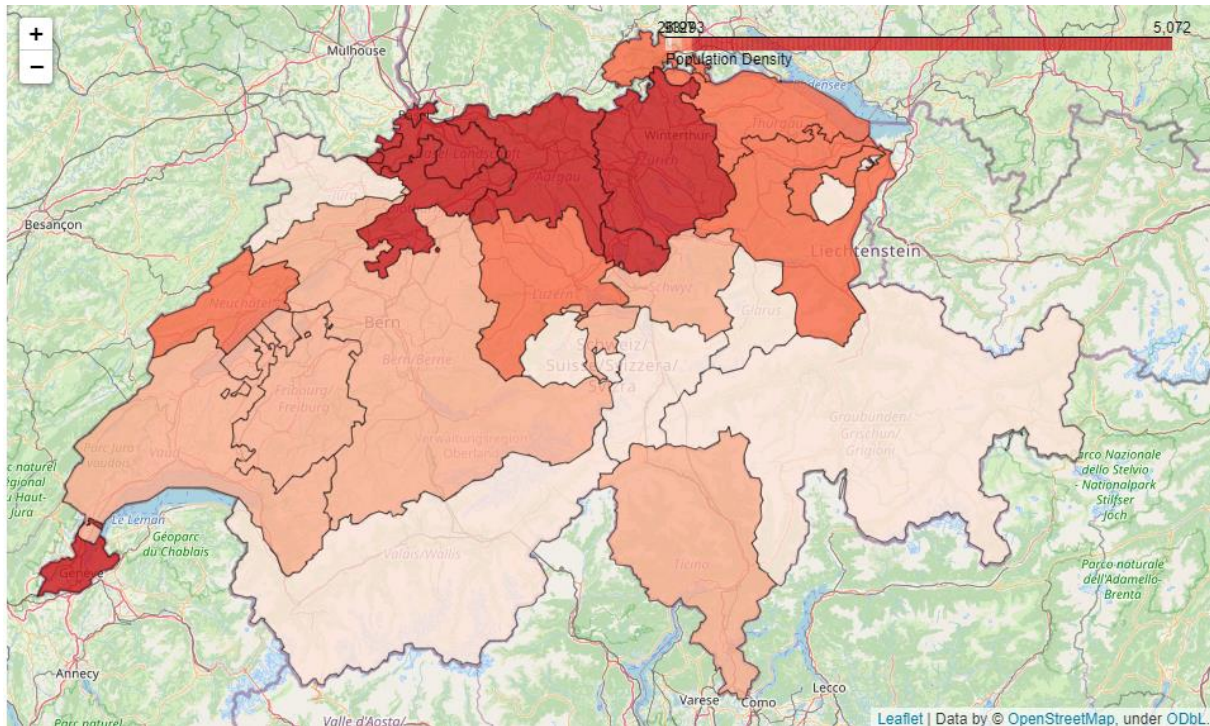
With the clustering, all data is available to be mapped and interpreted.

4. Results

Map 1: GDP per canton



Map 2: Population Density per canton

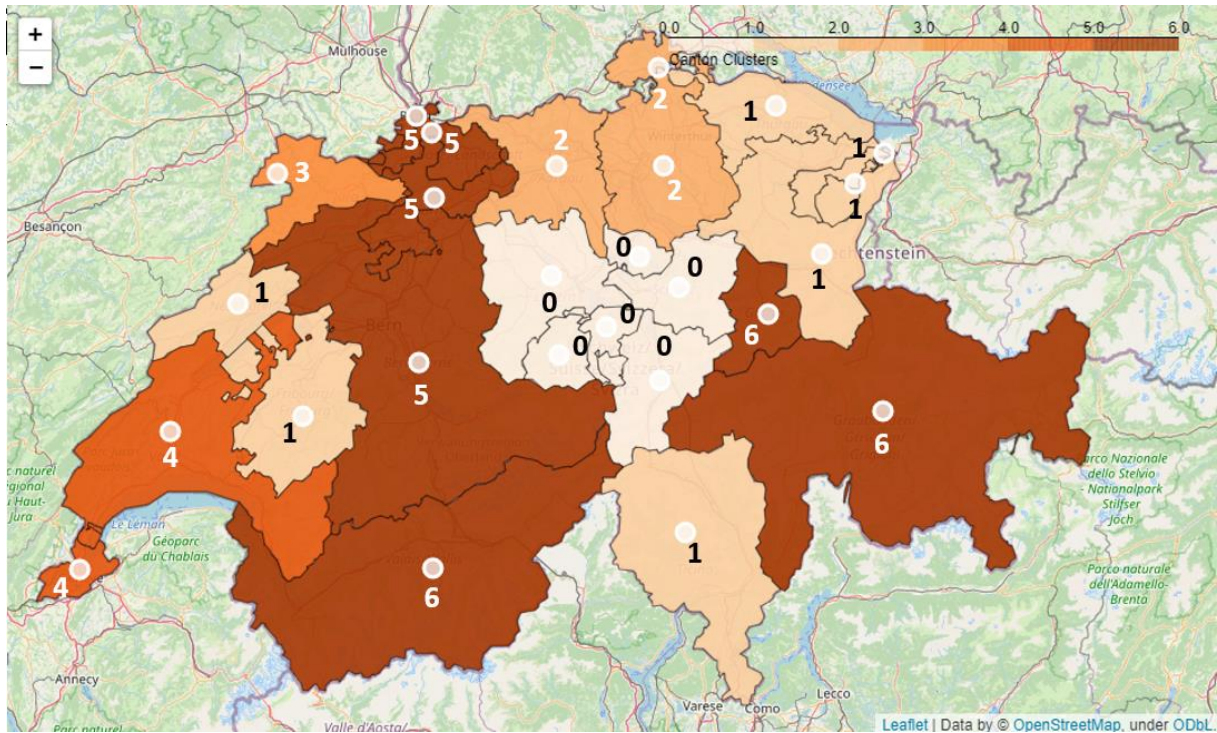


To map both GDP and population density, the data was binned along the quantiles

These two maps allow some interpretation, with we will further apply to the canton clustering:

- Population is most dense in the central northern region and in the south-west
- In the center, central south and east of the country, population density is lowest
- Additionally, there is one canton with low population density in the north-west. This is Canton de Jura, it catches the eye because it is surrounded by more densely populated cantons.
- GDP is highest in six cantons: Basel-Stadt, Canton de Genève, Canton de Neuchâtel, Ticino, Kanton Zug and Kanton Zürich.
- There is belt of low GDP cantons from the south to the north
- Canton de Jura, which is an island of low population density, also has rather low GDP and again is surrounded mostly by cantons with higher GDP.

Map 3: Clustering of cantons:



In the clustering map, pop-up markers are added to indicate the cluster label. In this map, in addition to the coloring, the clusters are indicated with numbers for easier readability of the map.

Cluster 0 is built by cantons in the center of the country, cluster 1 is more diversly distributed, cluster 2 consists of 3 cantons in a region with high population density, cluster 3 is Canton de Jura only, again an interesting region, cluster 4 is in the west, cluster 5 and 6 contain the three largest cantons.

For further interpretation of the data, the summary table is needed.

Summary Table:

	Canton	Cluster Labels	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
5	Kanton Luzern	0	Hotel	Mountain	Swiss Restaurant	Restaurant	Park	Shopping Mall	Café	Scenic Lookout	Train Station	Bridge
6	Uri	0	Hotel	Mountain	Ski Area	Scenic Lookout	Swiss Restaurant	Train Station	Trail	Bridge	Hotel Bar	Pool
7	Schwyz	0	Hotel	Pool	Park	Swiss Restaurant	Mountain	Restaurant	Ski Area	Scenic Lookout	Waterfront	Farmers Market
8	Obwalden	0	Hotel	Swiss Restaurant	Mountain	Train Station	Park	Café	Ski Area	Bridge	Historic Site	Scenic Lookout
9	Nidwalden	0	Hotel	Swiss Restaurant	Mountain	Ski Area	Bridge	Café	Scenic Lookout	Restaurant	Park	Train Station
11	Kanton Zug	0	Hotel	Park	Pool	Swiss Restaurant	Scenic Lookout	Gym / Fitness Center	Waterfront	Mountain	Plaza	Restaurant
12	Canton de Fribourg	1	Hotel	Swiss Restaurant	Café	French Restaurant	Bar	Pizza Place	Ski Area	Chocolate Shop	Restaurant	Plaza
17	Appenzell A. Rh.	1	Hotel	Swiss Restaurant	Restaurant	Café	Italian Restaurant	Austrian Restaurant	Scenic Lookout	Coffee Shop	Ski Area	German Restaurant
18	Appenzell I. Rh.	1	Hotel	Restaurant	Café	Swiss Restaurant	German Restaurant	Austrian Restaurant	Italian Restaurant	Bar	Coffee Shop	Scenic Lookout
19	Kanton St. Gallen	1	Swiss Restaurant	Hotel	Restaurant	Café	Ski Area	Bar	Mountain	Trail	Dessert Shop	Resort
22	Thurgau	1	Hotel	Swiss Restaurant	Café	German Restaurant	Bar	Ice Cream Shop	Palace	Restaurant	Italian Restaurant	Park
23	Tessin	1	Hotel	Italian Restaurant	Swiss Restaurant	Plaza	Café	Restaurant	Mountain	Scenic Lookout	Beach	Supermarket
26	Canton de Neuchâtel	1	Café	Hotel	Restaurant	Swiss Restaurant	Bar	Brewery	Music Venue	Beer Garden	Museum	Italian Restaurant
3	Kanton Zürich	2	Swiss Restaurant	Café	Pool	Park	Bakery	Hotel	Bar	Plaza	Japanese Restaurant	Vegetarian / Vegan Restaurant
16	Kanton Schaffhausen	2	Swiss Restaurant	Hotel	Café	Airport Lounge	Bakery	Bar	Park	Coffee Shop	Beach	Zoo Exhibit
21	Aargau	2	Park	Café	Bakery	Swiss Restaurant	Pool	Bar	Hotel	Plaza	Japanese Restaurant	Vegetarian / Vegan Restaurant
28	Canton de Jura	3	Train Station	Supermarket	Hotel	Café	Sporting Goods Shop	Swiss Restaurant	Park	Grocery Store	Brewery	Steakhouse
24	Waadt	4	French Restaurant	Park	Bar	Restaurant	Hotel	Beach	Art Museum	Plaza	Church	Gym
27	Canton de Genève	4	French Restaurant	Hotel	Park	Bar	Coffee Shop	Ice Cream Shop	Dessert Shop	Ski Area	Beach	Department Store
4	Kanton Bern	5	Café	Park	Hotel	Bar	Plaza	Swiss Restaurant	Train Station	Restaurant	Ice Cream Shop	Tram Station
13	Kanton Solothurn	5	Swiss Restaurant	Bar	Hotel	Café	Park	Art Museum	Italian Restaurant	Restaurant	Concert Hall	Plaza
14	Basel-Stadt	5	Park	Hotel	Art Museum	Café	Swiss Restaurant	Supermarket	Plaza	Bar	Concert Hall	Bakery
15	Basel-Landschaft	5	Park	Hotel	Café	Art Museum	Swiss Restaurant	Plaza	Supermarket	Bar	Scenic Lookout	Concert Hall
10	Glarus	6	Swiss Restaurant	Hotel	Ski Area	Restaurant	Mountain	Bar	Trail	Resort	Scenic Lookout	Rest Area
20	Graubünden	6	Hotel	Ski Area	Restaurant	Swiss Restaurant	Mountain	Scenic Lookout	Italian Restaurant	Ski Chalet	Ski Lodge	Town
25	Valais	6	Ski Area	Swiss Restaurant	Hotel	Restaurant	Café	Apres Ski Bar	Mountain	Train Station	Town	Tea Room

5. Discussion

The Business Problem to solve are the questions where to expect what in Switzerland as a tourist, and where to find good life and work as an immigrant.

Tourist view:

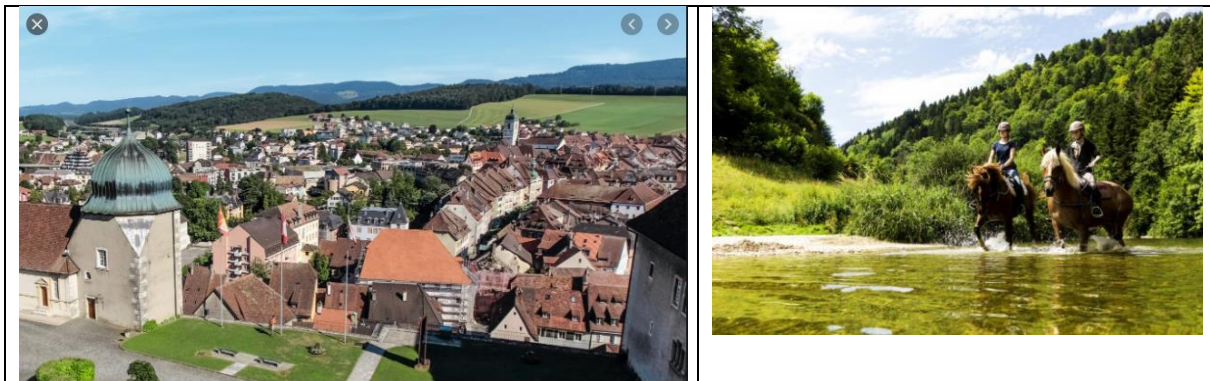
Cluster 0 contains many hotels, mountains, scenic lookouts and skiing areas. This certainly is an interesting area for a traveler.

Cluster 6 very prominently features skiing, mountains, resorts and scenic lookouts. At the same time, population density is low in this cluster. We can interpret that cluster 6 is an area to find high mountains, outdoor activity and scenery. Since hotels are prominently present in areas 0 and 6, we can expect to find many attractions in these clusters, but also many other visitors.

Within this cluster, it is also interesting to compare GDP. The two big cantons in there, Wallis and Graubünden, show low and intermediate GDP respectively. Following the hypotheses that both these cantons are mainly characterized by mountains (intact tourist infrastructure, prominent skiing facilities and low population density), the tourist sector seems to be more successful in Graubünden. This hypotheses could be verified or nullified by additional data mining and analysis.

Canton de Jura in the north-west has low population density, low GDP and does not feature tourist attractions like skiing areas, mountains etc. From the Top 10 venues, we can deduce that this is an area of low population and not many visitors. As a tourist seeking a Swiss vacation with nature, deceleration and little tourist infrastructure, this seems like the right place to go.

A quick image search on Google with 'Canton de Jura' as keyword seems to confirm this:



To live and work:

Most productive areas are Basel, Geneva, Neuchâtel, Ticino, Zug and Zurich. Within this cluster, the two financial centers Geneva and Zurich are found, along with the hotspot for pharmaceutical industry, Basel. In and around these cantons, population density is high. A glance into the top 10 venues in clusters 2, 4 and 5 shows places for daily life are prominent, so these are cantons where somebody seeking to immigrate into Switzerland would live with high probability, unless working in the tourist sector.

6. Conclusion

Only with three data elements, GDP, population density and venues, insights could be gained as to where in Switzerland to expect what. The following hypotheses were formulated:

- In the central cantons of Switzerland (cluster 0), tourist hot spots are found.
- For skiing and mountain experiences, the three cantons in cluster 6 offer opportunities
- Graubünden is the more successful tourist destination than Wallis.
- For the rural and decelerated experience, Canton de Jura should be considered more closely.

- Working opportunities are mostly found in the central north, north-west and south-west of the country, excluding opportunities in the tourist sector, for which cluster 0 and Graubünden should be considered.

To gain further insights, verify or nullify above hypotheses, further data could be analyzed, like housing prices, tax levels, university locations, types of economy etc.

Also, when displaying the seven clusters, coloring of the map could be improved to more clearly distinguish the clusters.