

Battle of Neighborhoods in Switzerland

Yueming HU

1 Introduction

1.1 Background

I live in Thun, a medium sized *commune* in the Canton of Bern in Switzerland. *Commune* here in Switzerland is similar to the concept of neighborhood.

Recently, I got a new job in another canton, i.e. Canton of Geneva, which has 45 neighborhoods. I need to choose a neighborhood to live.

1.2 Problem

What I'm satisfied most for my current neighborhood Thun is its number of food venues and the population size. Because I'm a foodie, so I hope there could be many different food-related venues, no matter it's Café, restaurant, snack shop or even tearoom. And I also like medium sized neighborhood, neither too crowded nor too quiet as no one in the street in the evening.

Therefore, I want to find a similar neighborhood as Thun in the new destination Canton of Geneva.

1.3 Target Audience

People who have to move to a new place and want to find a similar neighborhood as the one he or she currently lives in, with particular attention to food venues and population size.

2 Data Acquisition and Cleaning

2.1 Data sources

The websites <https://www.citypopulation.de/en/switzerland/geneve/> provides the names and population of each neighborhoods in Canton of Geneva, and the population of Thun is from Wikipedia (<https://en.wikipedia.org/wiki/Thun>).

Geopy: the geographical coordinates, i.e. latitude and longitude of all the neighborhood are obtained by using geopy.

Foursquare Api: the Food related information, including the name, category and number of food venues is obtained through Foursquare Api.

2.2 Data cleaning

Pandas is to read the tables in the above-mentioned website. There are 2 tables, I use index to call the table contains population data. Since I only need the latest population data, I extracted the column of poplation datated in 2019-12-31 and the name of each municipality, then create a dataframe as below.

.5]:

	Municipality	Population
0	Aire-la-Ville	1169
1	Anières	2387
2	Avully	1712
3	Avusy	1401
4	Bardonnex	2277
5	Bellevue	3380
6	Bernex	10244
7	Carouge	22621
8	Cartigny	986
9	Céigny	791
10	Chancy	1698
11	Chêne-Bougeries	12504
12	Chêne-Bourg	8674
13	Choulex	1188
14	Collex-Bossy	1709
15	Collonge-Bellerive	8299
16	Cologny	5598
17	Confignon	4631
18	Corsier (GE)	2186
19	Dardagny	1867
20	Genève [Geneva]	203951

The latitude and longitude of each neighborhood are obtained by using the geocode method of geopy.

Then the latitude and logitude data are used in the Foursquare requests to get relevant venue information. Since I want the information related to food, so I choose the category search to get the information. Accoding to [Foursquare](#), the category code of food is “4d4b7105d754a06374d81259”. The resulting json file contains many features, I extract the name and category of each food venue, as below.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Category
0	Aire-la-Ville	46.190592	6.042878	Café du Levant	Restaurant
1	Aire-la-Ville	46.190592	6.042878	Au Croissant Timbré	Bakery
2	Anières	46.276080	6.223412	Le Floris	French Restaurant
3	Anières	46.276080	6.223412	Le Cottage	Swiss Restaurant
4	Anières	46.276080	6.223412	Café F.	French Restaurant
...
399	Veyrier	46.166816	6.185681	La Réunion	French Restaurant
400	Veyrier	46.166816	6.185681	Boulangerie Briffaz	Bakery
401	Veyrier	46.166816	6.185681	Max Poulet	Food Truck
402	Veyrier	46.166816	6.185681	Pizza Da Guido	Pizza Place
403	Veyrier	46.166816	6.185681	Bersier	Restaurant

I use the same way to get all the names and categories of food venues in current neighborhood Thun.

2.3 Feature selection

When doing analysis with k-means clustering, the features I will use are the population number and the food venue number. So I use the groupby method to get the number of venues in each neighborhood.

The population number is converted into population in thousand, so that the feature values would be comparable and can be used directly. The plot of neighborhood features is as below.

