Capstone Project - The Battle of Neighborhoods

-----Analysis of the the most Venues around Aachen,Germany-----

INTRODUCTION

In week 3, we explored New York City and the city of Toronto and segmented and clustered their neighborhoods. Both cities are very diverse and are the financial capitals of their respective countries. One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are. But how will be this situation in a small city? Just like Aachen in Germany.

So if I want to look to open a restaurant in aachen, where should I open it? Similarly, if a contractor is trying to start their own business, where would you recommend that they setup their office?

DATA

In order to accomplish this goal, latitudes and longitudes of the Aachen are required. We can find the PLZ in Aachen and also where is assigned for these PLZs. We can use getcoder.arcgis('Aachen Germany') to get the location information of Aachen. From this data, I need to select the relevant shops and positions around the aachen and using their latitude and longitude values, I need to explore the venues with "art" section around Aachen by utilizing Foursquare API.

The data contains regions, corresponding latitude, longitude and country name along with many other features. I only use the features that I mentioned in this data and then filtered the regions.
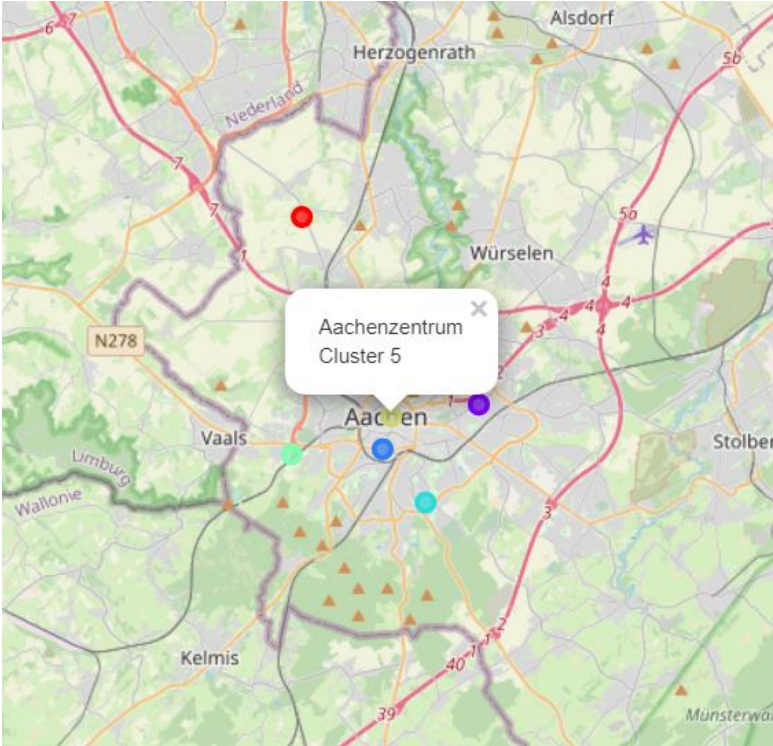
Then I used the acquired location data to explore the nearby art venues from the Foursquare API. Using the "explore" option, I look for top 10 art venues in 30 km radius for each city center and I get "Venue", "Venue Latitude", "Venue Longitude" and "Venue Category".
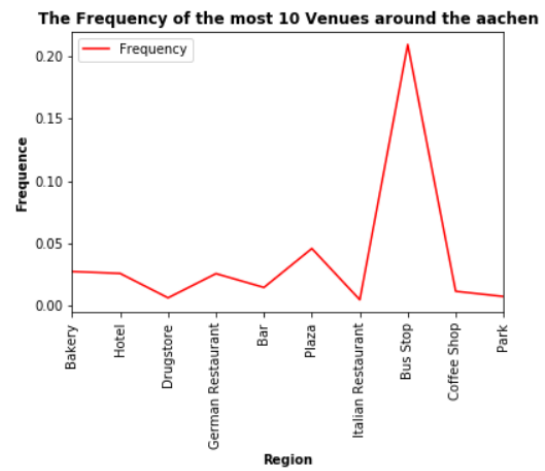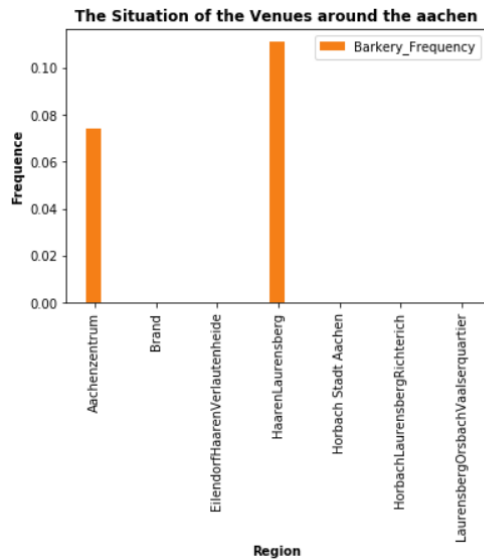
METHODOLOGY

Then we use the one hot encode and transform the data as we want. Finally we can get the relationship between the regions and the venues. Through the calculating the clusters, we set the clusters as 8, and then use the k-means algorithm and do the classification.

Results

We can get the cluster on the map and get the top 10 venues and further informations in Aachen and nearby, like the followed pictures.

	PLZ	Ortsname	Ortsteil	Landkreis	Bundesland	Latitude	Longitude	Cluster Labels	1st Most Common Venue
0	52062	Aachen	Aachenzentrum	Städteregion Aachen	Nordrhein-Westfalen	50.776410	6.085890	6.0	Bakery
1	52064	Aachen	Aachenzentrum	Städteregion Aachen	Nordrhein-Westfalen	50.768240	6.082041	6.0	Bakery
2	52066	Aachen	Aachenzentrum	Städteregion Aachen	Nordrhein-Westfalen	50.754500	6.099478	6.0	Bakery
3	52068	Aachen	Aachenzentrum	Städteregion Aachen	Nordrhein-Westfalen	50.779500	6.121089	6.0	Bakery
4	52070	Aachen	HaarenLaurensberg	Städteregion Aachen	Nordrhein-Westfalen	50.790965	6.095144	5.0	Hotel
5	52072	Aachen	HorbachLaurensbergRichterich	Städteregion Aachen	Nordrhein-Westfalen	50.826941	6.050155	0.0	Bus Stop
6	52074	Aachen	LaurensbergOrsbachVaalserquartier	Städteregion Aachen	Nordrhein-Westfalen	50.766860	6.045922	3.0	Flower Shop
7	52076	Aachen	BrandKornelimünsterKrauthausenLichtenbuschNüth...	Städteregion Aachen	Nordrhein-Westfalen	50.708715	6.165719	NaN	NaN
8	52078	Aachen	Brand	Städteregion Aachen	Nordrhein-Westfalen	50.756490	6.162285	2.0	Gastropub
9	52080	Aachen	EilendorfHaarenVerlautenheide	Städteregion Aachen	Nordrhein-Westfalen	50.786631	6.150915	4.0	Ice Cream Shop
10	52072	Aachen	Horbach Stadt Aachen	Städteregion Aachen	Nordrhein-Westfalen	50.826941	6.050155	0.0	Bus Stop
11	52076	Aachen	Friesenrath Stadt Aachen	Städteregion Aachen	Nordrhein-Westfalen	50.708715	6.165719	NaN	NaN

The Situation of the Venues around the aachen

The Frequency of the most 10 Venues around the aachen

DISCUSSION

By different k number will get different results. Maybe we should find the best k and improve the accuracy.

CONCLUSION

So, according to the data, the Bus stop is the most popular place, hhhh, but maybe it is useless. The most popular restaurant is pizza shop, An the 1st Most Common Venue is bakery in some regions.