Introduction/Business Problem

Zurich City, Switzerland becomes more and more a place for food with new restaurants opening every month. On the other hand, investment and operating costs for restaurants are very high. To be succesful, it is important to choose the right location. This report gives an overview of current restaurant offering by cuisine in the city's districts. Potential restaurant owner may use this information when considering locations to open new restaurants.

Data

We will use data about restaurants from Foursquare and evaluate them according to their location (district) and cuisine. In additionally we may use other data such as rent-prices for restaurant properties.

Methodology

We First look into the separation of districts and for that we choose the official statistical districts. We can get data incl. Geo json from the open data portal of the city of Zurich. In the same portal we also find data about officially registered restaurants per district.

In a next step we include data from foursquare to have additional data of restaurant types.

Finally we use KNN to cluster the different districts in the same way we did for New York, but with only Data about Restaurants.

Results

The first discovery we made is that when looking at the distribution of restaurants, we have a high density in the "central" districts whereas a very low density in most of the "outer" districts.

In a seconf step we use data from Foursquare to add more information about the restaurants. For this report, we focus on restaurant type, but we could also add other features. Unfortunately the amount of data from foursquare for this city/country is not that high.

We can cluster the district with KNN according to the restaurant types to see where we have similar types of restaurants.

Discussion

Foursquare data is probably not the best source for this area/country. In a follow up, we should try to use data from google maps which seems to have more data and thus would lead to more features and better significance.

Conclusion

We can see that there are some centers with a lot of restaurants and variotions of restaurant types. Then we have some districts with very few restaurants where it's not stastistically significant. We need to identify further data such as renting prices etc. to give a better guidance on our initital question.In terms of clustering,

we have identified similar districts according to similar restaurant types and could use this to e.g. build a recommender for which restaurant type that is still missing could be popular in which district.