Neighborhood Segmentation to find the best Location for a Game Restaurant

Introduction

While many young people are looking for healthy and sustainable alternatives to conventionally sourced meat there is still a source of meat that is actually underrated. Game often is found in regions where their natural predators are extinct. Forest owners are looking for ways to protect areas of reforestation which are as important as ever, regarding their role in reducing CO2 emissions. Also certain species contribute to the spread of dangerous viruses such as African Swine Fever and Influenza. Therefore it is evident that certain game populations need to be regulated. Hunting may have a bad reputation with some people in the society. Yet there is an increase in demand of game, because of the so to speak negative CO2 balance. While game is mostly found in rural areas, many cities suffer from increasing populations of e.g. raccoons and wild boars. In order to help hunters and land owners with the sales of game, and in order to meet the demands of an inceasingly ecologically thinking urban population the intent is to open a restaurant that provides meals using responsible meat sources, including game but also with meat from cattle from extensive farming when no other land use is possible.

Problem

How can we find the best location for the restaurant in order to have the right kind of people who care about the environment so that the restaurant can be run efficiently?

Approach

We can use Neighborhood Segmentation and the Machine Learning Algorithm k-means to come up with the kind of neighborhood suited to the opening of this restaurant. We are looking for venues attracting younger and alterniative people such as coffee shops, event locations, yoga studios among others.

Use of Data

For our restaurant we choose the city of Berlin, Germany. The wikipedia page for Berlin's district will give us the names of boroughs and districts using web scraping techniques such as beautifulsoup. In order to find latitude and longitude we will use Geocoder and create a dataframe that combines borough, neighborhood and the location. With these data the first folium map of Berlin can be created. Then, we add the Foursquare API derived data and can find venues for each neighborhood, e.g. the ten most common in each neighborhood. After using one-hot-encoding, the k-means algorithm will be applied and we can find the right cluster of neighborhoods for our restaurant