The Battle of the Neighbourhoods

Ahmed B

1. Introduction: Business Problem

Zurich, Switzerland is home to a large number of financial institutions and banking companies, and the low tax rates attract overseas companies to set up their headquarters there.It is also a great tourist destination and is home to a few UNESCO world heritage sites. In this project we will determine the optimal location to open a restaurant in the Zurich district, and will be targeted at stakeholders interested in opening a restaurant in Zurich. Since we are not aware of what type of restaurant is popular in Zurich, we will also determine the most popular type(s) of restaurants in Zurich. We will utilize data science and machine learning techniques in order to do this. This project will be of interest to any one wanting to open a restaurant in Zurich, and as a secondary application can also be useful to anyone wanting to visit or live in Zurich.

2. Data

In this project we will utilize three datasets in order to solve our problem. The first dataset is a list of municipalities and information regarding them such as which district it is in, the date of its merger and its coat of arms. The latter two columns were dropped as they are not needed for this project. This data set was scraped from Wikipedia. The next data set is one contain the latitude and longitude for each municipality. This dataset was obtaining from taking the coordinates from each of the Wikipedia pages of the municipalities and pasting them into a CSV file using excel. The file containing the list of venues was obtained using the Foursquare API. We selected all venues not just restaurants in order to see where restaurants were more popular.

Some of the features extracted from the data was the district of the municipality, the latitude and longitude of each municipality, and from the foursquare API we extracted the venues, the venue category, and its geo-coordinates. All of this data can be used to help us in our analysis and problem solving, which will be discussed more thoroughly in the methodology and analysis sections.