IBM Data Science Capstone Project

Picking right location for a new restaurant in berlin

Senura Fernando

November 2020

1. Introduction

The IBM Data Science Professional certificate course on Coursera concludes with a Capstone Project. This project is about using data science toolset on a real-life problem and demonstrating the creation of value by applying the learned skills. This report presents the capstone project which is done for picking right location for a new restaurant in Berlin.

2. Problem Definition

a. The Problem

Berlin is the capital and largest city of Germany by both area and population. In an competitive field, Berlin is making a strong claim to be the world's premier city. The German capital can lay claim to reasonable rent, an English-speaking population, thriving businesses, great food and nightlife, and a burgeoning startup scene.

In here, business problem is to picking up a best location to open a new restaurant in Berlin. Taking the population, price level at which the restaurant will operate. The intent is to find an optimal location in an area which is easily accessible for tourists and for wealthier local citizens as well.

b. Assumptions and business logic

The assumption behind the analysis is that , unsupervised machine learning can be used to create clusters of neighborhoods of Berlin that will give with a list of areas for consideration for the restaurant. The intent is that restaurant to be situated close to one of the gastronomical centers and touristic hotspots.

c. Audience

This analysis could be useful for group of market players and business people who have idea to open new restaurant in berlin and also tourists to take idea about restaurants around Berlin

d. Data

To perform this analysis , following data required:

1. List of the boroughs and neighborhoods of Berlin
2. Geo-coordinates of the boroughs in Berlin
3. Top venues of boroughs

List of boroughs will be obtained from Wikipedia.

(https://en.wikipedia.org/wiki/Boroughs_and_neighborhoods_of_Berlin)

Geo-coordinates of boroughs will be obtained with the help of the geocoder tool in the notebook.

Top venues data will be obtained from Foursquare through an API