COURSERA IBM DATA SCIENCE CAPSTONE PROJECT

Opening a new Azerbaijani-Turkish restaurant in Amsterdam, Netherlands

Introduction

In this project I will try to find an optimal location for an Azerbaijani restaurant in Amsterdam. Specifically, this report will be targeted to stakeholders interested in opening a Restaurant specialized in Azerbaijani cuisine in Amsterdam, Netherlands.

The idea behind this project is that there are already some number of similar restaurants in Amsterdam but adding to cuisine the Azerbaijani food will be the uniqueness of the brand-new place. Since Azerbaijani cuisine has been shaped by Persian, Caucasian and Turkish cultures, we will examine the last 2 categories of restaurants in the area due to current political and cultural connections. With the purpose in mind, finding the location to open such a restaurant is one of the most important decisions for this entrepreneur and I am designing this project to help him find the most suitable location.

Business

The success of a new restaurant depends on several factors: demand, brand loyalty, food quality, competition, and so on. In most cases, a restaurant's location plays an essential determinant for its success. Hence, it is advantageous and of utmost importance to determine the most strategic location for establishment in order to maximize business profits.

Data

To solve this problem, I will need below data:

- List of neighborhoods in Amsterdam, Netherland
- Latitude and Longitude of these neighborhoods
- Venue data related to restaurants in the area

Methodology

The neighborhoods data were found on kaggle.com - the world's largest data science community. Geographical coordinates for each neighbourhood were extracted from there too. As for Amsterdam's census data — median household income, total population, and population of Azerbaijani across the European Country — European Statistical System provides it. (I didn't use it for this specific project, but had a look kept in mind during my analysis)

For returning the number of restaurants in the vicinity of each neighbourhood, we will be utilizing Foursquare Developer API, more specifically, its explore function. One must register for a Foursquare developer account here to access their API credentials.

After getting the full data set I analyzed the numbers of restaurants in "Caucasian Restaurant", "Azerbaijani Restaurant" and "Turkish Restaurant" categories. Since the last one had more or less enough number of venues to analyze and see the trend.

Lastly, I performed the clustering method by using k-means clustering. K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. It is one of the simplest and popular unsupervised machine learning algorithms and it is highly suited for this project as well. I have clustered the neighborhoods in Amsterdam into 3 clusters based on their frequency of occurrence for “Turkish Restaurant”. Based on the results (the concentration of clusters), we can predict the ideal location to open the new restaurant.