Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Sushi Restaurant in Athens, Greece

By: Spyridon Milas

Introduction

Over the last few years the capital city of Greece, Athens has attracted many international kitchens with great response from the local customers as well as tourists. Dining areas developers are taking advantage of this trend, establishing brand new restaurants and hiring experienced chefs in order to correspond to this increasing demand. As a result, the location of a significant investment in a dining establishment is a serious consideration taking into account the competition that this industry is facing.

Business Problem

The objective of the capstone project is to analyze and afterwards select the best neighborhood locations in Athens, Greece to open a new sushi restaurant. Using Machine learning techniques like clustering, and Foursquare APIs with a combination of raw text data sets from open sources, this project aims to provide solution to the following question: If a food service developer would want to open a new Sushi Restaurant in Athens, which neighborhood would most likely provide the higher revenue income?

Data

To solve the business problem described above, we will make use of the following data:

- List of neighborhoods of Athens, capital of Greece. This will define the locational scope of the project.
- Latitude and longitude coordinates of these neighborhoods. These will be necessary to plot the city map and map the relevant venues.
- Venue data about sushi restaurants initially. Then venue data about every dining area within the most favorable cluster in order to end up to the resulting neighborhood with the fewest competition threat.

Sources of data

The Wikipedia page
(https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Athens) contains a list of neighborhoods in Athens, with a total of 63 neighborhoods. Web scrapping techniques will be used in order to extract that data from the Wikipedia page, and with the assistance of Python requests and beautifulsoup packages will be converted to an editable data frame. Consequently, using Python Geocoder package we will get the latitude and longitude geographical coordinates of the neighborhoods in order to be used by the Foursquare API to get the nearby venues.

Using the Foursquare API, we will get the necessary data for the venues around our neighborhoods. Initially, we will isolate our data results for Sushi Restaurants in order to form our necessary clusters. Afterwards, venue data concerning dining areas in general will be used in order to locate the neighborhood with the fewest restaurants in the cluster, we determined that it is the most appropriate. In the next section, we will further analyze the steps that we followed in order to get to the final result and the machine learning techniques that will be used.