Coursera Capstone

IBM Applied Data Science Capstone

Finding a Location for a New Italian Restaurant in Chicago, Illinois

Introduction:

Chicago is the largest metropolitan area in the Midwest. With its large population, the city presents prime opportunities for young entrepreneurs to launch brand new businesses such as restaurants or cafes. Opening a new restaurant in Chicago can possibly prove to be a good decision depending on if the restaurant itself can attract customers. For this report, I'm exploring where the best locations would be to start a new Italian restaurant in a Chicago neighborhood. Properly choosing the correct location to start the restaurant is critical for attaining long-term success and keeping the restaurant open. When selecting neighborhoods to start the restaurant, there are a few key variables you have to take into consideration. These variables include which neighborhoods show a high demand of restaurants and what neighborhoods have surrounding venues that can attract customers to the area.

Business Problem:

The primary goal of this project is to find the best possible neighborhoods to place a new Italian restaurant in Chicago. This will be done by using various methods previously learned in this course such as clustering and utilizing Foursquare location data. The main audience this project would apply to are entrepreneurs looking to start a new restaurant or current Italian restaurant owners who may be looking to relocate in the city of Chicago.

Data Description:

- Create a dataframe that contains all the neighborhoods in the city of Chicago. This dataframe will be created by scraping the data from the Wikipedia page (https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago). This page contains a list of all the neighborhoods in Chicago.
- Use the Python geocoder function to find the geographical points of every neighborhood in the city of Chicago. This will provide us with every neighborhood's respective longitude and latitude values. After finding the longitude and latitude values for each neighborhood, we will add these values to our dataframe.
- Create a folium map to visualize the neighborhoods of Chicago.
- We will then utilize Foursquare API location data. This will enable us to see information about venues in each neighborhood. We will use the Foursquare data to see what neighborhoods have Italian restaurants as one of their most visited venues.
- Use k-means clustering to determine the proper number of clusters that should be used to visualize the data. After we determine the number of clusters, create a folium map that visualizes the clusters. We will then be able to perform observations based on the results.