IBM Applied Data Science Capstone

Opening a New Pizzeria in Salt Lake City, UT, USA

By: Matt McDermott
January 2020

Introduction

Pizza. Undeniably one of the world's most popular dishes. It's versatility with the different pizza toppings to choose from, different methods of cooking it and different quality whether you order pizza at a high end Italian restaurant or have it delivered by your local pizzeria, one thing for certain is almost everyone loves pizza. In the United States, pizza was ranked first in favorite comfort food according to a recent survey conducted by The Harris Poll. In the same survey, 21% of people said if they had to eat one food for the rest of their lives, pizza would be that dish. Second place was chicken with 16% of the votes. It's clear that pizza is a favorite for many people, especially in the United States, but are there enough pizzerias? Salt Lake City for example is well below the average pizzeria per capita. Should there be more pizzerias in Salt Lake City, Utah?

Business Problem

The objective of this capstone project is to analyze where the best location in Salt Lake City, Utah would be if you wanted to open a new pizzeria. I will be using data science methodology and machine learning in order to find the best location for a new pizzeria.

Target Audience

This project will be useful for potential restaurant owners or investors in Salt Lake City. As mentioned before, Salt Lake City is below the national average for pizzerias per capita and since pizza is the nation's most popular food, it would be ideal to open more pizzerias in Salt Lake City.

Data

To solve the problem, I will use the following data:

- List of neighborhoods in Salt Lake City.
- Latitude and longitude coordinates of those neighborhoods in order to plot a map and also to get venue data.
- Venue data of other pizza restaurants in SLC. This will be used for clustering in the neighborhoods.

Data Source

This wiki page (https://en.wikipedia.org/wiki/List_of_Salt_Lake_City_neighborhoods) has the list of neighborhoods in SLC that I will extract using Python requests. Once extracted, I will use the geographical coordinates of the neighborhoods using Python Geocoder which will provide the latitude and longitude coordinates for the neighborhoods.

I will also use Foursquare API to get the venue data. With Foursquare's large database of 105+ million places and 125,00+ developers it will provide valuable venue data, especially for pizza restaurants which will help with our business problem stated above.