

IBM

Data Science Capstone Week 4

Introduction

This is a capstone project for the IBM Data Science professional Certificate. In this project, I will be testing my hypothesis that there are not enough German restaurants in the Toronto area. With this lack of German restaurants, this may provide a good opportunity for entrepreneurs to open new restaurants. I will be finding the most suitable location for a new restaurant to open.

Business Problem

The main goal of this project is to find the best location to open a new German restaurant in Toronto, Canada. I will be using data science methods and machine learning algorithms to provide an answer to this question.

Question: "Where should I open a new German restaurant in Toronto, Canada?"

Target Audience

Any person who wants to find the best place to open a German restaurant.

Data

We will need the below data to solve the problem:

1. List of neighborhoods in Toronto, Canada
2. Latitude and Longitude of the neighborhoods
3. Venue data related to German Restaurants

Extracting the Data

1. Scraping data of Toronto neighborhoods located in Wikipedia data set. This will be done with the imported library of BeautifulSoup4 which is a web scraping library.
2. Getting the latitude and longitude information of the neighborhoods using the Geocoder package.
3. Using Foursquare API to get venue data related to these neighborhoods.