Using Machine Learning to find locations to open a Peruvian Restaurant

1.Introduction

1.1 Background

For this Capstone project, I am creating a hypothetical scenario for a concept Peruvian restaurant who wants to explore opening an authentic Peruvian restaurant in Toronto area. The idea behind this project is that there may not be enough Peruvian restaurants in Toronto and it might present a great opportunity for this entrepreneur who is based in Canada. As Peruvian food is very similar to other Latin America cuisines, this entrepreneur is thinking of opening this restaurant in locations where Latin Amercican food is popular (many Latin America restaurants in the neighborhood). 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.

1.2 Business Problem

The objective of this capstone project is to find the most suitable location for the entrepreneur to open a new Peruvian restaurant in Toronto, Canada. By using data science methods and machine learning methods such as clustering, this project aims to provide solutions to answer the business question: In Toronto, if an entrepreneur wants to open a Peruvian restaurant, where should they consider opening it?

1.3 Target Audience

The entrepreneur who wants to find the location to open authentic Peruvian restaurant

2.Data

To solve this problem, I will need below data :

- List of neighborhoods in Toronto, Canada.
- Latitude and Longitude of these neighborhoods.
- Venue data related to Latin restaurants. This will help us find the neighborhoods that are most suitable to open a Peruvian restaurant.

3. Extracting Data

- Scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods