

Introduction

Opening and maintaining a restaurant is not an easy task. It requires proof of concept, dedicated staff, proper marketing, and strong financial understanding. To open a restaurant, you would first need to know the restaurant style and location. This project will first locate the top two restaurant venues in Toronto neighborhoods, then will locate a neighborhood that can open an Arabian sweet bakery.

Data sources

We will need a list of Toronto neighborhoods and their geographic locations. In addition, we require the Foursquare location data to provide the list of restaurants and ratings in each neighborhood. The list of Toronto neighborhoods can be found in Wikipedia, where we scraped the postal code, borough, and neighborhood [1]. The geographic locations can be downloaded from an online source [2].

Data Cleaning:

The neighborhood data first had to be parsed from Wikipedia into a data-frame indicating the postal code, borough, and neighborhood. After extracting the information into a data-frame the data-frame was checked to ensure the parsing step was successful and to remove any missing information. The list of neighborhoods and the geographic locations were merged together to give all neighborhoods in the data-frame geographic locations.

The foursquare location data provided a list of restaurants by inputting the data-frame created previously. The list of restaurants had to be cleaned to remove any category that is not helpful in the analysis stage such as cafeteria and restaurant. When the data cleaning stage is finished the received information had 89 neighborhoods and 88 restaurant categories.