

Open a Chinese Restaurant in Toronto

Xizhen Zhang

January 20, 2020

1. Introduction

1.1 Background

Toronto is the most populous city in Canada, and it's also a multicultural city. There are many ethnic neighborhoods such as Chinatown, Koreatown, little India, little Italy and so on, and it will be a significant factor to determine which neighborhood would be the best option to open a new Chinese restaurant. In addition, Ontario is one of the most immigrant-friendly province in North America, hence more and more Chinese are moving to Canada for their rest of life. Furthermore, Chinese food is a very popular food, and people all over the world like to eat it. Therefore, open a Chinese restaurant in Toronto probably will be a good choice to start one's business.

1.2 Problem

In this project, we will analyze the neighborhoods in Toronto to identify the most profitable and least competitive area. Then we can decide the location of the restaurant, where it can yield more profit to the owner.

1.3 Interest

This project focuses on three kinds of people:

- The person who interested in cooking and hope to open a Chinese restaurant as a side business. This analysis will tell them what are the pros and cons about this business.
- The business person who wants to get involved into Chinese restaurant field in Toronto. This analysis will help them to get enough information about the catering industry in Toronto.
- The person who does not have ability of cooking and needs to eat outside. This analysis will help them to know the distributions of Chinese restaurant.

2. Data Acquisition and Cleaning

2.1 Data Acquisition

- 1) Using “List of Postal code of Canada: M” table from Wikipedia, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, where can get the postal code, borough and neighborhoods in Toronto.
- 2) Using http://cocl.us/Geospatial_data csv file to get geographical coordinates for each postal code in Toronto.
- 3) Using statistic data from Wikipedia, https://en.m.wikipedia.org/wiki/Demographics_of_Toronto, where can get the distribution of population by their ethnicity in Toronto.
- 4) Using Foursquare’s explore API to get the present information about venues in Toronto.

2.2 Data Cleaning

- 1) In “List of Postal code of Canada: M” table, there are many not assigned values in column “Borough” and “Neighborhood”. First, I decided to scrape the rows that with “Neighborhood” not assigned, and if a cell has value in “Borough” but not assigned in “Neighborhood”, the “Neighborhood” will be the same as the “Borough”. As we can see, there are some neighborhoods exist in on postal code area, then I combined these two rows into one row and separated in “Neighborhood” by “,”.
- 2) Adding the geographical coordinates to existing table. First extracting the data from csv file, and then combining it with existing table by merging with same postal code.
- 3) Extracting the data from Wikipedia which give us the distribution of population in different neighborhoods in Toronto, and rename some columns to make us to understand easier.
- 4) Using Foursquare’s explore API to retrieve the information in Toronto, and it will return a JSON file. Then turn it into dataframe. After that, set the radius of area that we need to find.

The rest will be done in next week!

3. Exploratory Data Analysis

4. Predictive Modelling

5. Results and Discussion

6. Conclusion