

Finding a Suitable Sushi Restaurant in Toronto



Dezhen Kong

1. Introduction

In this project, I am aiming to create a hypothetical scenario for people who want to open a new sushi Restaurant in Toronto. Sushi has been a well-known Asia food for years and is popular in Toronto, therefore, it has a huge market in Toronto. The idea of this project is derived from there maybe not enough sushi restaurant in some areas of Toronto which cannot satisfy the need of public. By investigating the tendency of sushi restaurant distribution, this project serves as a guide for people want to open a sushi restaurant.

2. Data acquisition and cleaning

To solve this problem, the following data are used.

- List of neighborhoods in Toronto
- The geographical information (latitude and longitude) of these neighborhoods.
- Venue data related to sushi restaurant

2.1 Data acquisition

- Scrapping the list of neighborhoods in Toronto via Wikipedia.
- Getting latitude and longitude data via Geocoder package.
- Using Foursquare API to get venue data related to these neighborhoods.

2.2 Data cleaning

There are several problems in Neighborhoods data:

First, the neighborhoods data downloaded has a lot of borough that is “not assigned”. Therefore, I only use cells with assigned Borough in this project.

Second, more than one neighborhood can exist in one postal code area. Therefore, I merge two rows with the same postal code into one row with the neighborhoods separated with a comma.

Third, some cell has a borough but a not assigned neighborhood. In this project, these neighborhoods will be the same as the borough.

After fixing these problems, the neighborhood data is ready to merge with the geographical location data.