

Finding the best place for a restaurant in Toronto

Vladimir Volkov

19/10/19

1. Introduction

Toronto has an estimated population of just over 2.8 million in 2016, which makes it the 4th most populous city in North America the most populous Great Lakes city. Over the next 20 years, Toronto is expected to continue its moderate growth, surpassing 3 million by 2026, and reaching nearly 3.2 million in 2036.

Toronto is a global city filled with vast opportunity and is home to an array of distinctive and dynamic neighbourhoods that reflect the diversity of its population.

As the population of Toronto is constantly growing, this is a good place to open a restaurant business.

Usually restaurants are concentrated in developed areas.

The easiest way to find the best place for a restaurant is to search for the area where the largest number of restaurants are already concentrated. It will also be useful to identify the most popular types of restaurants Toronto.

2. Question

What types of restaurants are the most popular in Toronto? What neighborhood has maximum density of restaurants?

2. Data acquisition and cleaning

2.1 Data sources

I'll use Wikipedia page with the list of Neighborhoods in Toronto to create the dataframe with the list of Toronto neighborhoods. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

To obtain geographic coordinates, the file will be used at the following link: https://cocl.us/Geospatial_data

Forsquare base will be used to search for venues in Toronto neighborhoods, including restaurants.

2.2 Data cleaning

Scraped data from the Wikipedia page will be cleaned in order to remove "Not assigned" postal codes.

It is also needed to replace missing data in column "Neighbourhood".

Aggregated by postal code Data will be combined with geographic coordinates.

To get a list of restaurants, first I need to get a list of all the places in Toronto from the Forsquare database. Then this list will be filtered by category of the venue including string "restaurant" in its name.

The final dataframe for analysis will include following information: Neighborhood, Neighborhood Latitude, Neighborhood Longitude, Venue (Name of a restaurant), Venue Latitude, Venue Longitude and Venue Category (Type of a restaurant).