Finding the Optimal Neighborhood to Expand Restaurant Business

Malachi Jones

March 14, 2021

1. Introduction

A problem facing many successful restaurants is where they should put their new locations when they decide to expand their business. One way a restaurant can attempt to find the best area for their new locations is to search for the areas that are most like where their current restaurant is located. It makes sense to try and find similar areas because given that the initial restaurant was successful, a restaurant in an area with similar characteristics should have a higher probability of success.

In this report, we will be analyzing a specific scenario under these guidelines. It is assumed the original restaurant is in the Lawrence Park East neighborhood of Toronto, Canada. The restaurant owner wants to open a new location either in Toronto or another Canadian city. The question is what city and neighborhood has characteristics most like Lawrence Park East. While this report uses a specific example, the methodology and general framework can be applied for different neighborhoods and cities.

2. Discussion of Data

2.1 Data Sources

For the analysis, neighborhood location data and data on the types of venues in the area will be needed. Two data sets are needed to get the necessary information. First, a data set consisting of a list of all the postal codes in Canada along with the name associated with that postal code and its latitude and longitude coordinates. This data set came from AggData (https://www.aggdata.com/free/canada-postal-codes). To get data on the types of venues in each area, the Foursquare API is used.

2.2 Data Cleaning

First, the postal code dataset was downloaded and cleaned. After importing the data set, the postal codes from the desired cities were selected. For the analysis, only the top seven cities in Canada by population were used. It is likely that the areas most closely resembling the area with the original restaurant will be in an urban area. These cities are Toronto, Montreal, Vancouver, Calgary, Edmonton, Ottawa, and Winnipeg.

Next, venue data was imported using the Foursquare API for each of the postal codes. The venue type was extracted for each venue and grouped by postal code. For each postal code, the proportion of each venue type compared to the total venues in that area was calculated. For example, Calgary (Cranston) has a proportion of 0.036 for American restaurants. This means that 3.6% of the venues in the area are American restaurants. The proportions along with the names of each area were then put into a table. In total, there are 187 different areas in our data set. Also, there are 329 different categories of venues.