

Introduction

Background

Although the global food service market, defined as sale of food and beverages for immediate consumption, was worth US\$ 3.4 Trillion in 2018 and is expected to reach a value of US \$4.2 Trillion by 2024 [1], starting a new business in the restaurant industry nowadays is challenging. That's due to the existence of various competitive players in the market. In addition, there are large fast-food chains that expand easily from one location to another. According to the Foodservice Industry Forecast report in Canada, the commercial food service sales are expected to grow to US \$77.5 billion in 2020 [2].

Many small businesses start and fail everyday due to the lack of prior market surveys and studies to understand the possible needs for the targeted business locations. For example, in the food service market, critical factors can affect the success or failure of a business depending on the geospatial location of such a business as well as the existing rivals. Thus, a proper market study is required to understand the success factors of starting a business in the restaurant industry.

Problem

It is required to identify the optimum location(s) for starting a new restaurant in Toronto, Ontario in Canada. In this context, the optimum location can have several meanings. For example, it could mean a location where there is minimum competition, or a location where there is a weak competition or. These two options will be addressed in this report.

Proposed solution

To address the problem of identifying the optimum location for starting a new restaurant, the two suggested definitions of the optimum locations need to be defined as in the following:

- 1) A location with a minimum competition means a location where the number of restaurants is low with respect to the number of people or the current population living there.
- 2) A location with a weak competition means a location where there are some restaurants, but their quality or rating is low.

Data Acquisition and Cleaning

Data Sources

The type of data as well as the sources are listed in Table (1) below.

Table 1 - Data types and sources

Data type	Source
Postal Codes, Boroughs and Neighbourhoods in Canada	Wikipedia [3]
Postal Codes, and geographical locations in Toronto, Ontario	Coursera (CognitiveClass.ai)

The population in each Neighbourhood in Canada in 2016	Canada Statistics [4]
The geographical boundaries for each Neighbourhood in Toronto, Ontario	Canada Statistics [4]
Venues' IDs, names, locations, distances from the corresponding Neighbourhood centre location, postal codes and categories type	Foursquare API
Restaurants' ratings, likes, checkins, price tier and tips	Foursquare API

The first two data sources are used to identify the geographical location and postal code of each neighbourhood in Toronto. The 2016 population data from Canada Statistics is then linked to these two data sources using the postal code in both data sets. The geographical boundaries data area used to facilitate the plotting of data in choropleth maps. Using the geographical location of each neighbourhood, Foursquare API is used to get all the restaurants with their details such as ratings, like count, check-ins, tip count and price tier.

It should be noted that the population data available online is that of 2016. So, this data has been used current data from Foursquare API due to unavailability of population data in 2020.

In the Foursquare API, all the venues in each neighbourhood were searched using the intent keyword “browse” and with a limit of “1000” and a radius of “1000” m. This resulted in all the venues within a 1 km circle radius in each neighbourhood. The number of restaurants were more than 500 which is the maximum number of premium calls that can be made with Foursquare API. So, it took more than one day to get the details for each restaurant.

Data Cleaning

From the Wikipedia table of postal codes that starts with “M”, all the unassigned boroughs have been removed, and all the unassigned neighbourhoods have been assigned to their corresponding boroughs. The total number of neighbourhoods was 103.

In the 2016 population data, the postal codes that starts with “M” were only chosen. However, one postal code (namely “M7R”) didn’t have a population data so it was removed. The total neighbourhoods with given population data became 102.

The geographical boundaries were obtained from Canada Stats website in the form of a .shp file. The procedure of converting this file into .geojson file using QGIS[5] is well explained here [6].

Using the Foursquare API, all the venues for each neighbourhood was searched. Then, only the venues that had the keyword “Restaurant” in the venue category column were chosen. Since some neighbourhoods were close to each other (or to be specific, they were closer than 1000 m which was the radius indicated in the search query for venues using the Foursquare API), there

were duplicates in the restaurants. They were not many though. All the duplicates were removed. Moreover, not all the neighbourhoods had venues. At the end, there were 881 restaurants in 99 neighbourhoods in 9 boroughs.

References

- [1] <https://www.prnewswire.com/news-releases/global-food-service-market-report-2019-2024-market-is-expected-to-reach-a-value-of-us-4-2-trillion-300907559.html>
- [2] <https://www.restaurantscanada.org/resources/foodservice-industry-forecast/#preview>
- [3] https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- [4] <https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm>
- [5] <http://qgis.com>
- [6] <https://medium.com/dataexplorations/generating-geojson-file-for-toronto-fsas-9b478a059f04>