The Best Location for Vegetarian Restaurant in Toronto

1. Introduction

1.1. Background Information

Canadians are going meatless. In September 2018, Charlebois, Somogyi, and Music from Dalhousie University conducted a study, which consisted of an online survey administered to Canadian consumers to determine their attachment to meat consumption and willingness to adopt a plant base diet. The survey was carried out in 3 days and they gathered 1027 samples. Based on this study, Charlebois, Somogyi, and Music estimated that over 6.4 million Canadians prefer to either reduce or eliminate meat consumption. The reasons for Canadians going meatless were found to be related to Health, Animal Welfare, the Environment and Taste.

The study revealed that the typical vegetarian or "to-be-vegetarian" is a woman, who is less than 38 years old, is highly educated, and her income is more than $75,000. Even though, health benefits gained by being a vegetarian seems to be important for both genders, women appear to be more concerned about animal welfare than men. Men, regardless of their age, consider eating meat as a great pleasure in life. Moreover, the less educated men love meat in general.

Here is the link for the Dalhousie study for further information.

1.2. Problem

As Canadians are going meatless, the new restaurants are needed that provide vegetarian meals. Toronto is a very diverse city and full of restaurants, café shops, and fast food places. However, vegetarian restaurants are hard to find, and the goal of this study of "The Best Location for Vegetarian Restaurant in Toronto" is to answer the question: what is the best neighbourhood to open a new vegetarian restaurant? This study uses the profile of the typical vegetarian provided by the Dalhousie study as a base line. First, this study tries to find out the neighbourhoods, where the typical vegetarians live. Once the

neighbourhoods have been found, the next step is to find out the existing venues in those neighbourhoods and determine the potential neighbourhood for establishing a vegetarian restaurant.

1.3. Interest

Establishing a new vegetarian restaurant or even many of them, is a great business opportunity for an individual(s), who has a spirit of an entrepreneur and may even have the same concerns (e.g. health, animal welfare, environment, taste) as the consumers in the Dalhousie study. Therefore, the end result of this study is an enormous help when figuring out what's the best neighbourhood for a vegetarian restaurant is.

2. Data Acquisition and cleaning

2.1. Data Sources

In addition to the information of the Dalhousie study, this project uses two datasets to find out where the potential vegetarians live. City of Toronto uses the census data of population in Canada to compile their own datasets to help government and community organizations with local planning. These datasets are open for everyone and can be found at City of Toronto's Open Data portal. This project is utilizing two of those datasets:

1. **Neighbourhoods**

   The City of Toronto has divided Toronto in 140 social planning neighbourhoods and they can be found in the Neighbourhoods dataset.

2. **Neighbourhood Profiles**

   City of Toronto has compiled the demographic, social and economic characteristics of the people and households in each City of Toronto neighbourhood of those 140 neighbourhoods and they can be found at the Neighbourhood Profiles dataset.

Also, the Foursquare is utilized to pull the list of venues near those neighbourhoods.

2.2. Feature Selection

The Neighbourhoods dataset includes 16 features and 140 rows, one row for each neighbourhood. The features needed from this dataset are AREA_SHORT_CODE, AREA_NAME, LONGITUDE, and LATITUDE. The AREA_SHORT_CODE can be used when matching neighbourhoods' data with neighbourhoods' profile data. The AREA_NAME, LONGITUDE, and LATITUDE are needed in order to the Foursquare to pull the list of venues near the neighborhoods.

The Neighbourhood Profiles dataset includes 163 features and 2383 rows. The first row has "Neighbourhood Number" which can be used when matching data with the neighbourhoods' data as it is the same as the AREA_SHORT_CODE in the neighbourhoods' data. This study is interested in data related to the number of populations that matches with the vegetarian profile in Dalhousie study (e. g. gender, age, income, education). The data that needs to be extracted are stored under the features of CATEGORY, TOPIC, DATA SOURCE, CHARACTERICS, and all those 140 neighbourhoods, which each has an own column. The purpose is to find out in which neighbourhoods the potential vegetarians and future customers live.

2.3. Data cleaning

The end result is one combined dataset that has only the information needed. The Neighbourhoods dataset didn't have any missing values or inaccurate data. From this dataset, the unnecessary columns were dropped. The Neighbourhood Profiles dataset required more work. It had lots of unnecessary data, which were not needed in this study. Several columns were dropped and only the rows that were necessary in order to find out vegetarian profiles and their locations were extracted. The dataset had missing data, which were filled in zeros. Also, most of the datatypes were converted from object to int64 to allow calculations. Finally, the cleaned dataset was transposed and combined with the Neighbourhoods dataset.