

IBM Applied Data Science Capstone Project

Restaurant Business Analysis In Toronto

Introduction

According to restaurantscanada.org, 2020 would witness commercial foodservice sales improvement by 4%. Alberta and Ontario will lead the way with 4.4% and 4.2% growth, respectively. By 2021, foodservice sales are forecasted to go over 100-billion dollar.

The restaurant industry represents 4% of Canada GDP, with around 85 Billion dollars in annual sales that are generated by the restaurant industry, it employs 1.2 Million people not less than 7% of the country total workforce. These figures could seem a bit staggering but when we put them in the context of Canadians making, daily, 22 Million visits to restaurants, it may completely change our perception of the restaurant business outlook.

Toronto, being the capital of Ontario, and having by far the biggest population and strongest economy within the province. Being a cosmopolitan metropole, it fosters diverse culinary backgrounds when it comes to food businesses, catering and, restaurants and attracting a huge number of immigrants makes its restaurant business a blooming one.

Business Problem

This project objective is to come up with an analysis of the restaurant business in Toronto based on several features or characteristics such as population levels, income ranges, restaurant business categories, geolocation within the city and so on. In other words, data science methodology will allow to answer the following question: Within the city of Toronto, where would be the best location to invest in a restaurant business and also for a given area within the city, what would best category of restaurant to invest on.

Target Audience

The target audience of this project are potential restaurant business investors or restaurant managers who need to conduct a market research in order to get specific insights and be able to determine the business arguments to invest in a given category of food establishment within a specific area with a unique set of characteristics. This project being specific to the city of Toronto aims to provide a more targeted view than reports or studies related to Canada and would help anyone interested in investing in this sector within Toronto.

Data acquisition

1. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M: this Wikipedia page provides the exhaustive list of Toronto postal codes or FSA, with their corresponding boroughs and neighborhoods.
1. "https://cocl.us/Geospatial_data" this CSV file provides the longitude and latitudes of the Toronto FSA or postal codes.
2. <https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm> - This repository from Statistics Canada, provides a boundary file representing the FSAs of Canada, this file is in Shape format. A conversion is to be conducted to transform this file to GeoJSON format using QGIS software - please note that a reduction of the file has been conducted to keep only Toronto FSA with the GeoJSON file.
3. <https://www.canada.ca/en/revenue-agency/programs/about-canada-revenue-agency-cra/income-statistics-gst-hst-statistics/individual-tax-statistics-fsa/individual-tax-statistics-fsa-2017-edition-2015-tax-year.html#t> this page of the Government of Canada, provides the Individual Tax Statistics by Forward Sortation Area (FSA) – 2017 Edition, this CSV dataset will provide us with income inputs, this will be one of the features of our analysis and machine learning algorithm.
4. <https://www12.statcan.gc.ca/> Statistics Canada provides a CSV file of the 2016 population census; the population level feature will be assessed to see whether it has an incidence on the restaurant business in Toronto.
5. <https://developer.foursquare.com/docs/api/endpoints>: Restaurant businesses category related data is retrieved via the Foursquare API; this category is identified within Foursquare database with the following ID 4d4b7105d754a06374d81259 (<https://developer.foursquare.com/docs>). An HTTP request will be sent to the API as follows: GET <https://api.foursquare.com/v2/venues/search>, it will return a list of venues, matching our category ID, near the specified location (providing its latitude and longitude).