A CARIBBEAN RESTAURANT IN MONTREAL

1. INTRODUCTION

1.1. Business Problem

Montreal is one of the major cities in Canada and Quebec in particular. Its population is the second largest French-speaking population in the world after France, composed largely of immigrants. One of the most immigrant-friendly cities in North America with more than half of the entire Haitian Canadian population residing in Montreal it is one of the best places to start a Caribbean restaurant.

In this project we will go through step by step process to make a decision whether it is a good idea to open a Caribbean restaurant. We analyze the neighborhoods in Montreal to identify the most profitable area since the success of the restaurant depends on the people, environment, ambience and other. Since we already know that Montreal shelter a greater number of Haitian than any other city in Canada, it is a good idea to start the restaurant here, but we just need to make sure whether it is a profitable idea or not. If so, where we can place it, so it yields more profit to the owner.

1.2. Interest or Target audience

Who will be more interested in this project? What type of clients or a group of people would be benefitted?

1. Business personnel who wants to invest or open a Caribbean restaurant in Montreal. This analysis will be a comprehensive guide to start or expand restaurants targeting the Caribbean crowd.
2. Freelancer who loves to have their own restaurant as a side business. This analysis will give an idea, how beneficial it is to open a restaurant and what are the pros and cons of this business.
3. Haitian or Caribbean crowd who wants to find neighborhoods with lots of option for Caribbean restaurants.
4. Business Analyst or Data Scientists, who wish to analyze the neighborhoods of Montreal using Exploratory Data Analysis and other statistical & machine learning techniques to obtain all the necessary data, perform some operations on it and, finally be able to tell a story out of it.

2. DATA ACQUISITION AND CLEANING

2.1. Data acquisition

A. We use the "Open Data" on *http://donnees.ville.montreal.qc.ca/dataset/quartiers* for data on the reference boroughs, neighbourhoods and cities of Montreal. For the geospatial coordinates of the neighbourhoods, we use the geolocator(geocoder). This data includes geospatial information (latitude and longitude). Neighbourhoods are reference neighbourhoods in housing.

B. The data on Montreal's immigrant population comes from the 2016 Census through the document entitled in French: "ANNUAIRE STATISTIQUE DE L'AGGLOMERATION DE MONTREAL / MONTREAL AGGLOMERATION STATISTICAL YEARBOOK". The document is available here: *http://ville.montreal.qc.ca/portal/page?_pageid=6897,68087755&_dad=portal&_schema=PORTAL*

C. To get location and other information about various venues in Montreal I'm using Foursquare's API. Using the Foursquare's explore API (which gives venues recommendations), I'm fetching details about the venues up present in Montreal and collected their names, categories and locations (latitude and longitude).
From Foursquare API (*https://developer.foursquare.com/docs*), I retrieved the following for each venue:

- Name: The name of the venue.
- Category: The category type as defined by the API.
- Latitude: The latitude value of the venue.
- Longitude: The longitude value of the venue.

2.2. Data Cleaning

After importing some libraries such as numpy, pandas, geocoder matplotlib and others, we proceeded to read the first csv file containing information on the boroughs of reference of Montreal. We must specify that the databases of the different sites of Montreal are in French and the names of these neighborhoods or boroughs are also in French. We have only changed the titles of the columns in English. The number of reference districts is 90.
The latitude and longitude of the districts are obtained from Geocoder (geolocator). A CSV file has been created for this purpose. Once finished with these preparations, we concat these dataframes.
As for the demography database in Montreal, as for the others, we have modified, adapted or translated the columns to finally obtain the distribution of immigrants of Haitian origin. It should also be noted that the data available in open data contains only the data for the districts and not for the neighbourhoods as we wished. We should make do with that.

Finished this first round, it's time to find the categories of the different venues from Montreal. To do so, we first proceeded to the geospatial location of Montreal, then from our Foursquare identifier locate in an area with a radius of 1000 meter, we get the JSON file of the categories and Nearby Venues from around 1174. This work resulted in 211 unique categories in which Caribbean Restaurant is one of them. Then follows the One hot encode for getting dummies of venue category. So that we calculate mean of all venue groupby their neighborhoods.

An example of what features can be extracted from the data by using Foursquare API:

	name	categories	lat	lng
0	Duc de Lorraine	Bakery	45.494618	-73.618279
1	Les Appartments Rockhill Inc	Residential Building (Apartment / Condo)	45.495024	-73.614498
2	Belvédère Summit Circle	Scenic Lookout	45.491326	-73.604463
3	La Caverne	Russian Restaurant	45.495546	-73.620538
4	Parc Summit / Summit Park	Park	45.491478	-73.604490