Capstone Project: The Battle of Neighborhoods — High-End Haitian Restaurant

Kenson Gilles Just now · 3 min read

Coursera Capstone Project — IBM Data Science Professional

As a Data Scientist equipped with the skills and the tools to use location data to explore a geographical location, we will have the opportunity to be as creative as we want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of our choice or to come up with a problem that we can use the Foursquare location data to solve.

Business Problem

Opening a new restaurant has always been every chef's dream, whether you are new or experienced in the culinary field you're both driven by the same charismatic law which is **passion**. Despite cultural differences between cuisines, customers still enjoy and appreciate each flavor set on their tables, nothing brings people together like a good dish. However, such a dream comes with a lot of planning, challenges, and let's not get started on the cost. It can be costly, time-consuming starting a business in the wrong neighborhood, or where restaurants are over-saturated.

This is where we come in as a Data Scientist, experts in data analysis. The objective of this project is to analyze and find the best location in Toronto Canada to open a new High-End **Haitian Restaurant**. Using Data Science tools and methodology we'll be able to determine where in the city would be the safest and most profitable neighborhood to open a Caribbean Restaurant. We will take a close look at each neighborhood in the city by analyzing the most common **venues**, **population**, and **criminal data** rate to help us determine the safest borough.

Data

Requirements

1. The list of Toronto's neighborhoods.

2. Toronto's neighborhood coordinates (Latitude & Longitude).

3. Toronto's crime rate data.

4. Neighborhoods venues data.

Sources

1. We'll use Canada's postal codes list found on Wikipedia.

2. Foursquare API for restaurant data.

3. Geocoder for the coordinates.

4. Due to limited data available, we'll use Wikipedia Crime in Canada dataset to analyze the criminal history.

Methodology

1. Collect, explore, and process the neighborhood data into the DataFrame.

```
# Let's import our python packages
import pandas as pd
import numpy as np
import requests

# We'll fetch and clean Canada's postal code data
---
postalDf.head(10)
```

	Postal Code	Borough	Neighbourhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront
5	M6A	North York	Lawrence Manor, Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
8	M9A	Etobicoke	Islington Avenue, Humber Valley Village
9	M1B	Scarborough	Malvern, Rouge
11	M3B	North York	Don Mills

12	M4B	East York	Parkview Hill, Woodbine Gardens
13	M5B	Downtown Toronto	Garden District, Ryerson

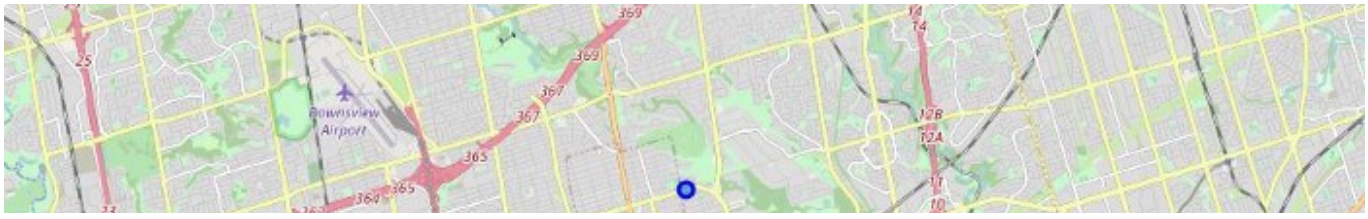
2. Assign the coordinates to each city accordingly using geocoder.

```
# Using geocoder we'll assign the city coordinates
!pip install geocoder
---
postal_merge.head(10)
```

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937

```
# Let's install and import folium to display our data on a map
!pip install folium
---
import folium
---

# create a map of Toronto using latitude and longitude values
map = folium.Map(location=[latitude, longitude], zoom_start=10)
```



3. Using FourSquare API retrieve all venues and filter the data received by anything related to Haitian or Caribbean restaurants.

```
# Using our FourSquare API credentials
CLIENT_ID = '<foursquare_client_id>'
CLIENT_SECRET = '<foursquare_client_secret>'
ACCESS_TOKEN = '<foursquare_access_token>'
VERSION = '20180604'
LIMIT = 30
```

4. Using our result we'll each city's crime rates to determine the safest neighborhood.

5. Using both data will be able to determine the most profitable and safest neighborhood in Canada.

6. Visualize the data on a map using folium.

—

We will continue to explore the possibilities of opening a profitable Haitian Restaurant. Stay tuned!