The Battle of Neighborhoods: Starting a Coffee Shop Business

Build models for segmenting the neighborhoods to find the most conducive locations for starting a business in Toronto City

A. Introduction

1. Business Problem

Kopiasli

AUTHENTIC COFFEE TASTE

Kopiasli (fictitious) is a coffee shop brand renowned for bringing a unique taste of authentic Indonesian coffee to your regular menus, such as Mocca Java, Toraja, Bali Blue, and even Luwak White Coffee. After successfully opening the first shop in New York City, **the cafe plans to expand its business in Toronto City**.

Toronto shares the same multicultural trait due to its large population of immigrants and is good for business. Thus, **the project goal is to figure out the city's most conducive locations for opening up a new coffee shop.**

To ensure this project's success, **the stakeholders require insights into the demographics, neighboring businesses,** and **crime rates**. For example:

- the cafes present in each neighborhood,
- the most popular cafes,
- the foot traffic of our target demographic,
- and the neighborhoods' crime.

Knowing that the project will have many things to consider, this project is open for discussion and targeted to the entrepreneurs, stakeholders, and/or data science enthusiasts who might encounter the same challenge and wonder how data science could be applied to the questions at hand.

2. Data Description

The followings are the data needed:

- **1st Data:**
  The most updated record of traffic signal vehicle and pedestrian volumes in Toronto City. The data is typically collected between 7:30 a.m. and 6:00 p.m at intersections where there are traffic signals.

- **2nd Data**:
  The most updated record of crime incidents reported in Toronto City provided by Toronto Police Services.

- **3rd Data**:
  The list of Toronto neighborhoods represented by postal codes and their boroughs. We will be using the Geocoder python package to retrieve the postal code's coordinates.

- **4th Data**:
  The popular or most common venues of a given neighborhood in Toronto. This information is stored inside Foursquare Location Data, and we will use Foursquare API to access it.

To sum up, we will use the 1st and 2nd data to analyze the pedestrian/vehicle volume and crime rates. Then, we load the 3rd data to obtain the exact coordinates for each neighborhood based on the postal code, allowing us to explore and map the city. Finally, we will use the coordinates and Foursquare credentials to access the 3rd data source through its API and retrieve the popular venues along with their details, especially for coffee shops.