Capstone Project - The Battle of the Neighborhoods

Diego Albuquerque

1. **Introduction\Bussines problem**

   Background

   The company XZY, have the interest to open a new restaurant Toronto, the speciality of company is regular restaurants always near a Parks, in this bussiness model the restaurants have the operation hours,menu, events and promotions based in the commom use of Parks, it is a specitility of the company.

   Seen this, we have the necessity of use data aproach to base the decision of in what place put a new restaurant.

   Problem

   Based in this context, we hav the question: **Where is the best place to open a new restaurante in Toronto?**, for this answer is necessary to find a place near a park and that not have many restaurants arround, in the vision of stakeholders, how less competition, best!.

2. **Data**

   To find the better place we have many option, drive a car arround the city finding places, use a helicopter or drone to fly above the city or the most easily and cheap is use the foursquare data do see the parks and restaurants arround then.

   In this way, we will map all the parks of Toronto, after all the restaurants and find the parks with less restaurants arround.

3. Libraries Which are Used

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution and Venues

XML: In Webscrap to separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module