Predicting the location for a new hostel

1. Background

Travelling is one of the most amazing experience anyone can have. The travel industry is one of the largest service industries in the world, an increasingly important one in the modern age. In the last decade traveling for tourism has been on the rise among the youth. Therefore, finding an affordable accommodation in a lively neighborhood is extremely important for anyone visiting a new country for a short stay. That's why hostels now a days have been taking over the accommodation scene.

2. Problem description

In this report we will be studying the neighborhoods of Toronto city in Canada to find the best location for a new hostel. The challenge here is to find the geographical data for the different venues in each neighborhood in the city, then start analyzing this data to find the locations characterized by having venues that are most important for any traveler visiting a new country for a short stay, such as, coffee shops, restaurants, bars, rental car services, etc.. within a walking distance from the target location. It is also important to reduce competition, so we should aim at locations with less number of hotels.

3. Data

To accomplish this analysis, I'm going to use the following data as follows:

   a. I will scrap the neighborhoods postal codes information from Wikipedia.
   b. Using this postal codes, I will then use geopy library to get the geographical coordinates of each postal code.
   c. Then using Foursquare API, I will gather the venues information within a radius of 500 meters around each location.
   d. After transforming this information to a data frame, it will then be used to classify the neighborhoods into clusters using sklearn library.
   e. The clusters with the most suitable characteristics for the location of a new hostel is going to be mapped to the neighborhoods with least number of hotels.