Finding Neighborhoods in Toronto to start an Indian Restaurant

Introduction/Business Problem

The main objective of this study is to make a list of ideal Neighbourhoods in Toronto to start an Indian Restaurant, based on the location data extracted using Foursquare API, and then their visualization through Folium. Toronto, being the most populous city and a major financial hub of Canada, attracts residents from all over the world both via employment opportunities and tourism. With nearly half, 46%, of its population coming from other countries. More than 200 different ethnic groups call Toronto home, and they bring with them over 140 languages. It is also one of those cities housing a wide selection of restaurants from cuisines all over the world, as we are going to see in this study.

Here, we deal with a hypothetical situation, of solving queries of someone planning to start an Indian restaurant in Toronto. They have approached us, to conduct a preliminary study of feasible neighborhoods for this new venture. Here, we use the business concept of proximity to competition, to shortlist ideal neighborhoods. Competition can be good, in industries where comparison shopping is popular. (That's why competing retail businesses, such as fast-food restaurants, antique shops and clothing stores tend to cluster together.) You may also catch the overflow from existing businesses, or customers who wish to try a different version of the same product, in our case, a new cuisine of food. Hence, in this project, we try to find the neighborhoods around Toronto, which has a clustering of Restaurants, but none of them being an Indian Restaurant. This will avoid competing with the same cuisine, one that's hard to pull off with a new restaurant.

The course of this project would involve accessing the data of different venues existing within a set radius from each neighborhood in the list. The collection of data about the different kinds of venues would be done using **Foursquare api**, and then only the categories containing the word "Restaurant" will be extracted from them. After this, we remove the data of Indian Restaurants from the dataframe, and then aggregate the other restaurant values together, to represent the relative presence of restaurants near each neighborhood. Then we add the Indian restaurant column back into the dataframe beside the aggregate value of restaurants of all other cuisines for comparison. Finally, we sort the dataframe in the decreasing order of aggregate restaurant value, and then pick the neighborhoods at the top, with a 0 value for Indian restaurants column