

Data Description

In order to analyze Toronto's neighborhoods, we need specific data that we can obtain from Foursquare.com.

More specifically, we need information about location and demographics that can help us understand which would be the neighborhoods that will increase the likelihood of success in opening a new restaurant.

For this objective, we are going to analyze population, income and second language information to identify the most populated neighborhoods, to assess if this highly populated neighborhoods have enough average income to increase the probability of success for a new restaurant, and the second language information will be useful to understand the ethnical and cultural characteristics of each neighborhood to try to match them with the kind of restaurant that can have potential in the identified neighborhoods.

We are going to use the k-means clustering technique to map, segment and cluster the neighborhoods with similarities in the analyzed characteristics and group them and understand those similarities to help us make the best decision in opening a restaurant.

With all this information we will make a recommendation on opening a specific kind of restaurant in a specific neighborhood or group of neighborhoods with objective information in order to increase its probability of success.