Description of the data and how it will be used to solve the problem

The data that are relevant for answering the business question is the location data of popular restaurants as well as the category of the restaurants. Before we use that data, however, we must determine what neighbourhoods or areas in Toronto that is suitable for opening up restaurants. It can be assumed since Toronto is the capital city of Canada; there are many restaurants in the whole city. However, the distribution of restaurants will not be uniform; there will be areas with more restaurants and areas that have fewer restaurants, even none. It is a simple assumption that the number of restaurants is proportional to the population density. The area with the highest population density then is the best location for opening up a new restaurant.

The city of Toronto is divided into several levels of areas: community council areas, wards, and neighbourhoods. The community council areas are the first level of area division. It consists of several wards, which in turn consists of several neighbourhoods. There are four community council areas, North York, Etobicoke-York, Toronto-East York, and Scarborough with different population densities. Toronto-East York has the highest population density compared to the other three according to the 2016 Canada census data. Since more population density means more potential customers, this community council area is selected to be the area for the new restaurant.

Toronto-East York is divided into several postal code areas and boroughs. Wikipedia provided the postal code data in Toronto, as well the corresponding boroughs and neighbourhoods. This data will be used to divide the community council areas further. The postal code data, combined with the latitude and longitude of each postal code area as provided by <http://cocl.us/Geospatial_data>, will provide a centre from which to search for popular restaurants. This data is finally going to be fed into the Foursquare search API to search for popular restaurants around the coordinates.

The resulting restaurant names from the Foursquare search API will be counted so that we will know which neighbourhoods or postal code areas that has many restaurants. On the assumption that if an area has many restaurants then that location is a good location to open up new restaurants, the location with the highest or high number of operating restaurants will be chosen as a good location for the client’s new restaurant site. This answers the first part of the business question.

The categories of restaurants from the Foursquare search API will also be counted and sorted according to the order of frequency and several of the high-frequency categories will then be chosen as the popular types of restaurants. This assumes that the number of certain restaurant category in an area is proportional to its popularity. This answers the second part of the business question.