**Introduction**

When opening a business, there are range of aspects to be considered. Among those, selecting the location to invest is of a paramount decision. This analysis is going to be a demonstration of how data science can help an investor for identifying a suitable place for opening a business in a particular city.

This analysis is focused on specifically locating a suitable place for opening a bakery in the city of Toronto, Canada. Location data from Foursquare API are the main source of input and on top of it the analysis is going to be elaborated. Finally, we will be exposing a list of neighbourhoods in Toronto where there are Bakeries reside inside the top 10 attractions within that neighbourhood. So that, it is advised to locate a neighbourhood excluding the ones in that list when locating a place for investing for the new bakery.

**Data**

This analysis is powered by the location data from the Foursquare API. Further, for illustration purposes folium maps are used. Apart from them, for data preparation, and for every aspect of the analysis, pandas library is used. Clustering from sklearn library has been employed for identifying prominent attractions inside a neighbourhood by categorizing neighbourhoods based on the features of the nearby venues.

**Methodology**First, the postal codes and boroughs in Toronto, Canada were scraped into the application together with the neighbourhoods for each postal district. Secondly, the latitude and longitude for each postal district were taken from a csv. file downloaded from coursera.

Then with the aid of the Foursquare API, the nearby venues for each neighbourhood in each postal district were retrieved together with ‘Venue Category’. The feature ‘Venue Category’ is important because it makes easy for us to identify the type of each nearby venue.

Then, the most common venue for each neighbourhood was identified through an analysis combining Foursquare data and scraped data that was mentioned above.

Subsequently, a clustering was carried out to identify groups with similar type of nearby venues. From those groups, it was able to identify the neighbourhoods where bakeries are there in the top 10 most common places.

Finally based on those clusters and the previous results, it was able to come up with a list of neighbourhoods that are better to exclude when deciding a place for the bakery investment.

**Results**

The analysis resulted a list of neighbourhoods that are advised to be excluded when deciding a place to open a bakery because, those neighbourhoods are already having popular bakery locations and opening a new bakery would reduce the return of the investment.

**Discussion**

The study recommends exclude the neighbourhoods in the following list in order to enhance the return of the investment while avoiding competitors. Because following neighbourhoods already have popular bakeries located. Below if the list of neighbourhoods that the study recommends to avoid,

Golden Mile, Clairlea, Oakridge, Wexford, Maryvale, Fairview, Henry Farm, Oriole, Studio District, St. James Town, Cabbagetown, Regent Park, Harbourfront, St. James Town, Berczy Park, Richmond, Adelaide, King, University of Toronto, Harbord, Kensington Market, Chinatown, Grange Park, Glencairn, Dufferin, Dovercourt Village, High Park, The Junction South, New Toronto, Mimico South, Humber Bay Shores, Mimico NW, The Queensway West, South of Bloor, Kingsway Park South West, Royal York South West, North Park, Maple Leaf Park, Upwood Park, West Deane Park, Princess Gardens, Martin Grove, Islington, Cloverdale.

**Conclusion**

This analysis was carried out with the hope of demonstrating how data science can aid for decision making when deciding a location to open a business. This particular study focuses on opening a bakery in Toronto, Canada while avoiding neighbourhoods having popular bakery locations already. Foursquare location data, folium maps and clustering were used as sources of information and python libraries towards the analysis. Finally the study was able to come up with a list of neighbourhoods to avoid when deciding the suitable location.