**Introduction/Business Problem**

Usually, an entrepreneur will open a branch for his company or shop if the first one is successful. For example, a restaurant owner which restaurant has many visitors will open branches for restaurants in other areas so that his income increases and his restaurant can be more widely known. The problem is, it is not easy to find a new good location to open the new restaurant. The owner tends to open the restaurant with the similar environment as the previous one and also consider the competitor in the new location.

If they want to open the branch in the same city or country, maybe they still can review the place by themselves. However, if they want to open the branch in another country, it will become more troublesome. For example, in this case is How do coffee shop owners in New York find the right location to open a branch in Toronto?

In this capstone project, I will try to use data science to overcome this problem. I would like to use the data to find the best location to open a new branch of the coffee shop in another country. This project will find out where is the similar location with the previous one and also the most common venue in that location to make sure the coffee shop doesn't have many competitors in the new place.

**Data**

In this project, I will be using one of the coffee shops in New York, that is For Five Coffee Shop, and then the data location in Toronto. I use the coffee shops data to find its latitude and longitude location, then find the environment (10 most common venues) on that location. Then, I use the Toronto neighbourhood data location to find 10 most common venues in each neighbourhood. After that, I will cluster the Toronto neighbourhood data with the coffee shop data and see in which cluster is the coffee shop will enter. After find the cluster, I will find the location which doesn’t have café or coffees shop to find the best location.