**Data Science Capstone – Comparing Restaurant Diversity in Toronto and New York**

# Introduction

Toronto and New York are the largest cities with their respective countries and are the most populated with old and new generations of immigrants. Therefore, we should expect a great diversity of restaurants around the cities. This project will discover the different restaurants in both cities and find the differences and similarities between them. This subject may be relevant to an aspiring group of chefs and businessmen willing to open a new restaurant in a North American city. The data insights generated may help them with choosing the best location to open, based on the type of food/cuisine that they offer.

# Data

The data used to solve this project are the location data, mainly the latitude and longitude, of the postal codes around both cities. This data is crucial in obtaining the restaurant related data with Foursquare Places API, which was used to search for restaurants around the cities.

# Methodology

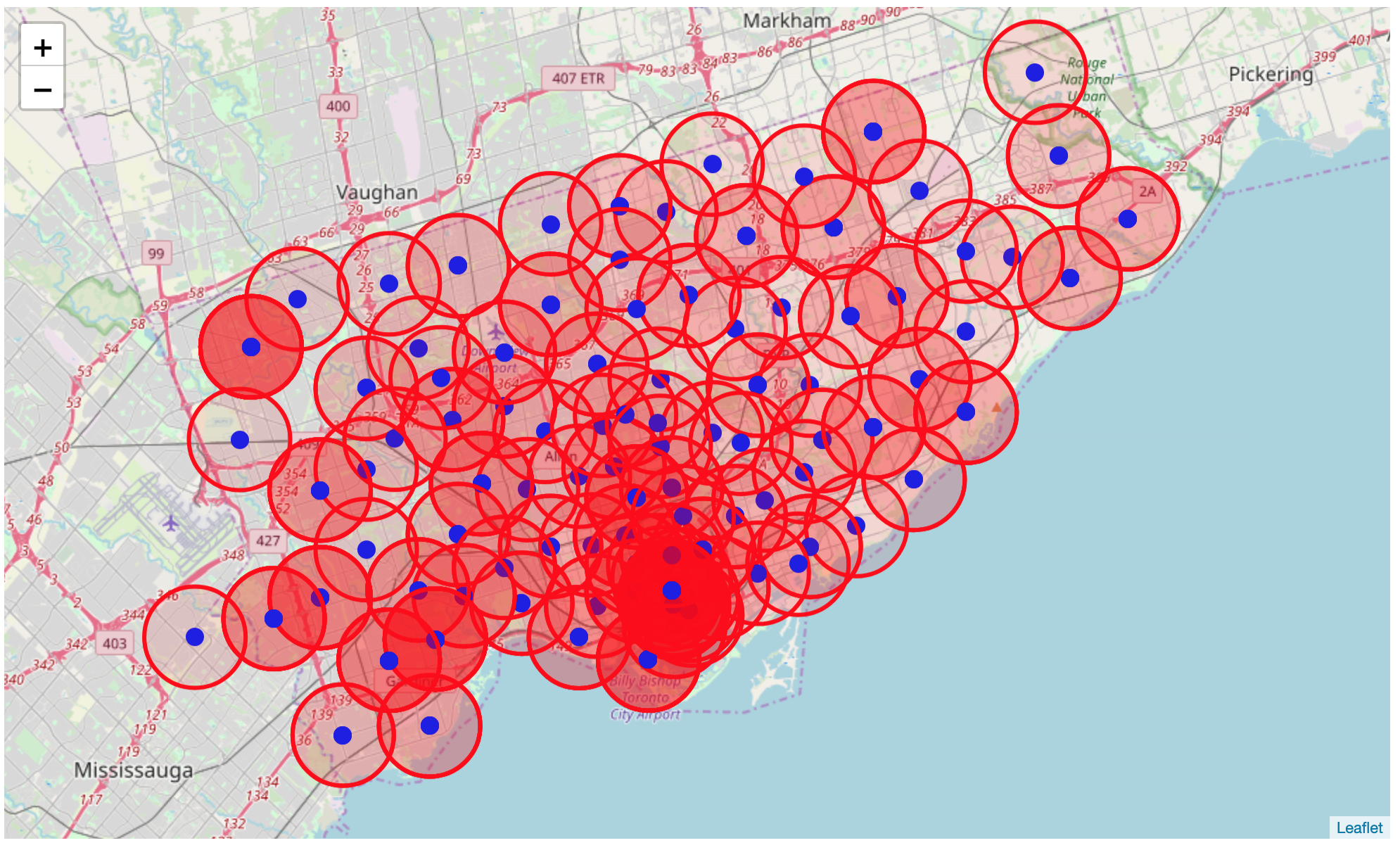
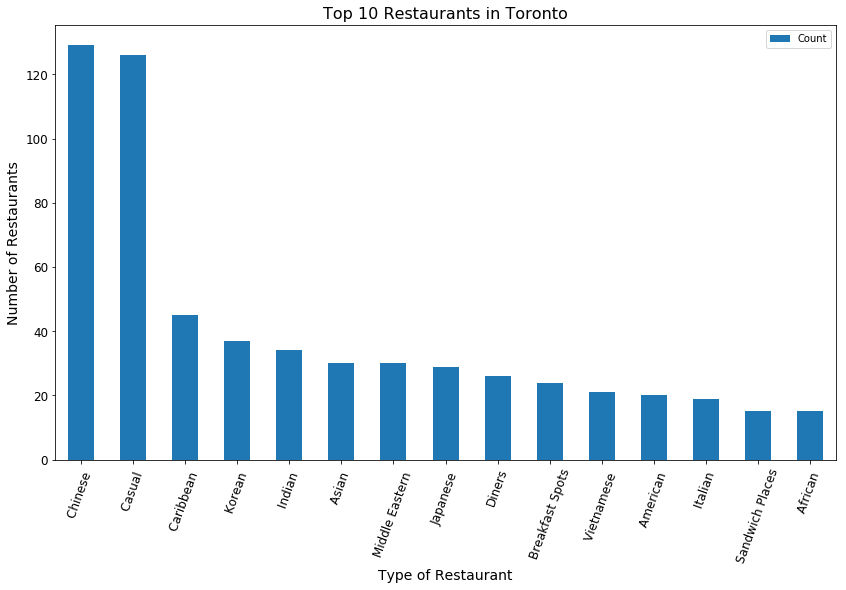
The Foursquare API was the main source of information used to obtain and collect the restaurant data for our analysis. In order to do that, one must specify a location, along with a few other optional parameters, to specify the types of venues that are of interest. In this case, a search radius was defined as well as the query ‘restaurant’. The postal code locations around the cities were used as centre points for the searching areas. In order to ensure that effective searching around the cities was possible, the searching areas were plotted with Folium, and an iterative process of selecting the radius was done until the most suitable radius was found for both cities where sufficient coverage was possible.

Figure 1: Searching Area Coverage in Toronto

All the location search results were collected in single Pandas dataframe for every city. It should be noted that while many of the searching zones overlap that many duplicate searches will be show up, however, pandas can easily get rid of duplicate data.

The restaurant data was then wrangled from its original format and separated into the names of the restaurants, category, address and location. From there, the category data was explored along with their respective restaurants and re-labelled if necessary, all the restaurants without a category were placed with the appropriate label.

# Results

The category labels were grouped and sorted for both cities; they were also visualized as bar charts as shown below.

# 

# Discussion

The graphs of both Toronto and New York show the major similarities and differences right away. First and foremost are the numbers of Chinese restaurants as being the most abundant significantly compared to other ethnic restaurants. Casual restaurants that are not associated with a specific cuisine rank as second for both cities followed by other types of restaurants. Starting from third place is where the real comparison begins and presents the different cultures of the cities. It is quite known about the special relationship between pizza and New York, which evident above along with Italian restaurants. Caribbean restaurants seem to also be popular for both cities. New York city also shows an abundance of Mexican and Latin American Restaurants which Toronto seems to lack. Toronto however is more abundant with Middle Eastern restaurants which are not as popular in New York.

# Conclusion

Toronto and New York are the largest and the most culturally diverse cities belonging to Canada and the US. Restaurants and cuisine in general are one of best ways for people to express their ethnicities and culture, especially within those two melting pots of cities. The analysis does in fact showcase the different cuisines that both places have to offer, with Chinese restaurants being most abundant. The analysis also gives us a general viewpoint of the different ethnic minorities in both cities.