Coursera IBM Data Science Specialization Capstone Project

'Determining the best location to open a café in Toronto'

1. Introduction

A person wants to open a cafe in the city of Toronto. Seemingly a straightforward decision, Toronto is a big city with a vibrant culture as well as financial hub of Canada and thus has intense competition for opening and running a business. It is also imperative for any business owner that the investment made reaps a decent level of returns. It is thus advantageous that the decision of deciding a place for opening the cafe happens by utilising the right data. We will be building a clustering model to sort the locations in Toronto and then recommend a place to open a café.

The results of the same analysis can also be used by any other business owner. Since the model will segregate the location, we can optimally decide where he/she should set up the new business to maximise possible success of the business.

2. Data

The data required for the model will be imported from the Wikipedia page of Toronto giving us the neighbourhood and borough names and the location data along with the nearby places will be queried using Foursquare API. The location data becomes especially crucial in the case of our problem statement, that of opening a café. The possible footfall of the café may be highly influenced by a lot of factors such as proximity to office spaces/residential places/educational hotspots or being in an upscale locality. This data can be readily explored using the API and will help us reach a solution. Moreover, the Foursquare API will provide us data regarding the locations of existing cafes and that will also help in deciding the ideal location for opening a new one.