**Report   
Applied Data Science Capstone Project**

Restaurants are a difficult business to operate as they involve huge risks. They often require huge capitals and resources, are an employee-intensive business and very often offer low to moderate margins alongside some cut-throat competition. There are more dining establishments and more diners now than during any other time in history; the National Restaurant Association reports over 1 million restaurant locations in the U.S. alone, a particular boon considering more than half the American population visits at least one of them each week. To fill the void and remain relevant, prospective, as well as established, restaurants have to hedge their bets with well-rounded and well-directed dataset. Data science provides valuable insights regarding market trends and evolving consumer lifestyles and tastes so that restaurateurs can better address the public demand and cope up with their ever-changing requirements.

**Business Problem**

People’s tastes are not attuned to a single restaurant. Some people like calm and tranquil places to dine in; others prefer crowded ones with sundry diversity. Some visit with their loved ones, others with their families, while some may opt for friends or colleagues. The preferences are as myriad as the possible offerings, with specific generational cohorts preferring one thing, “people who like Indian food” preferring another and everyone else liking a million other things in between.

My project user, a successful Indian restaurant chain owner in India and Canada is looking to expand operation into North America through New York (NYC).

**Problem Statement**

To locate and recommend to my user which neighbourhood in New York City will be the best choice to start a restaurant.

**Data Sources**

Before opening a new restaurant in NYC, my project will provide data and insight to:-

* Part 1: NYC Population & Demographic characteristics

Data source: <https://en.wikipedia.org/wiki/New_York_City> ; <https://en.wikipedia.org/wiki/Demographics_of_New_York_City>. Web scraping techniques have been used to get NYC's population density and demographics data from Wikipedia.

* Part 2: Who are the competitors in that location?

Data source: <https://cocl.us/new_york_dataset> and Foursquare API. These dataset was used to explore various neighbourhoods and each Indian restaurants venues in the neighbourhood.