**Coursera Capstone**

IBM Applied Data Science Capstone

# Opening a new restaurant in Colombo, Sri Lanka

*By: Ashane Fernando*

*December 2019*

# Introduction

## Background

Colombo is the de-facto - though no longer the official - capital of Sri Lanka. Nevertheless, it is still a burgeoning commercial hub with most of the economic activity in Sri Lanka being centered in this region. Over the last couple of years, there has been a dramatic rise in foreign investment in Colombo, due to the Sri Lankan government's initiative to transform Colombo into a "South Asian Las Vegas"

As a consequence, this has resulted in a rapid boom in tourism in Sri Lanka which in turn has triggered a meteoric rise in commercial development, with many restaurants, entertainment venues such as nightclubs, and hotels popping up around Colombo almost overnight. While Sri Lanka has always been a popular tourist destination, in the past it was mainly due to scenic beauty and natural attractions, and thus was mainly focused in the rural areas outside of the capital. However, at present, there is a large influx of tourists who come to Colombo solely to embrace the flourishing nightlife and sample a wide variety of cuisines from all over the globe.

While there are already many restaurants in Colombo which attract a substantial clientele, they are not sufficient to cater to the number of foreign tourists flooding into the city, in addition to the locals who already patronize these establishments. Thus, there is ample opportunity for a new restaurant to thrive, provided that they position themselves well; both in terms of location as well as business strategy.

## Business Problem

There is a saying among marketing professionals that three things are important in business: location, location and location, and when it comes to selecting a location for a restaurant this still holds true. Even within the city of Colombo, not all neighborhoods are the same, and the location you select for your restaurant can determine its a success or failure.

This project aims to apply the data science methodology, data analysis, and machine learning techniques to address the business question; if someone were to open a new restaurant within Colombo, where would you recommend that they open it?

**Target Audience**

This project is particularly geared towards local or foreign investors or restauranteurs who would be interested in opening or investing in new restaurants in the Colombo area.

For instance, earlier this year, in recognition of the growth of the restaurant sector in Sri Lanka, several of the leading restauranteurs in Sri Lanka banded together to launch the Colombo City Restaurant Collective (CCRC) in order to act as a shared voice for restauranteurs in Colombo and to facilitate coordination and collaboration with government authorities to guide the development of the restaurant industry[1].

Thus, this project may be of interest to current restauranteurs as well, who may wish to expand further.

# Data

When considering the location for a restaurant, we need to consider the other venues that surround it. In addition to the presence of other restaurants, we should also pay attention to the number of hotels and entertainment venues that are in close proximity. A large concentration of hotels would mean that there are more tourists in the area who are likely to visit your restaurant. In the same way, entertainment venues, such as bars, nightclubs and theatres would also attract more people to the area which once again translates to more potential business for your restaurant.

Thus, our primary focus would be on gathering information on the hotels, restaurants and entertainment venues within the neighborhoods of Colombo.

## Data requirement

In order to address this problem, we need to collect the following data:

- The list of neighborhoods within the Colombo city limits which we could use to divide the city into distinct locales.
- Latitude and longitude values which we could use to define the locations of each of these neighborhoods.
- Venue data for each of the neighborhoods, with particular emphasis on the hotels, restaurants, and entertainment venues. We need to obtain the location coordinates as well as the type of venue.

## Data Sources and Data Collection

Sri Lanka is a small country, with a land area of only 65,000km, and therefore it is not as easy to find structured geographical and location data as it would be in the case of larger and more developed countries. Nevertheless, there are still many tools we could leverage to obtain the information we need.

The Wikipedia page https://en.wikipedia.org/wiki/Postal_codes_in_Sri_Lanka contains the list of post codes for the whole of Sri Lanka. We can use web scraping techniques, such as the Beautiful Soup Python library or directly use the Pandas library to read the html directly from the website in order to obtain the list of neighborhoods in Colombo.

 We can then use the Python Geocoder package, which is free to use, to obtain the latitude and longitude coordinates of each of these neighborhoods.

Next, we can use the Foursquare Developer API to obtain the venue data for each of the neighborhoods. Foursquare is one of the largest location data platforms in the world and boasts a database for over 105 million places. Its API is used by many other technology giants such as Uber and AirBnB. We can leverage the Foursquare API ourselves (which is available for free, with some limitations) to gain information on the venues present in each of the platforms, including but not limited to their geographical coordinates and the type of venue.

# References

[1]       http://www.ft.lk/business/Leading-restaurateurs-unite-to-serve-up-Colombo-City-Restaurant-Collective/34-686048