



# IBM Data Science Capstone Project

*by Nicolas Dinh*

## I. Introduction

---

This project concludes of the IBM Data Science Certification Specialization. This goal of this last module is about to apply the knowledge and skills acquired during the course to a real-life problem. This project will be about solving a hypothetical business problem with real data (geolocation data from Foursquare using RESTful API calls as well as data scrapped directly from the web using libraries like BeautifulSoup) and analyzing it using Python in a Jupyter Notebook.

Similar to the approach that was taken during the course with New York and Toronto, the focus of this project will be about finding an optimal neighborhood (suburb) in the city of Melbourne in order to open a French restaurant.

## 2. Business Problem

---

Melbourne is a city with a very multicultural population without over 200 nationalities and over 230 languages spoken. The city is the nest of a large amounts of communities and hosts numerous cultural events from music, international festivals to shows, expositions and galleries. And one of the big results of this multicultural aspect is the diversity of restaurants and food that the city has to offer.

My client, a French chef and entrepreneur asked me to take a data-driven approach to find a good place to open a French restaurant. Ideally, it would close to the city center and in a popular suburb outside of the CBD that already offer a large choice of restaurants, bars and cultural events/gathering.

### 3. Data gathering

---

In order to solve the above business problem, the following data would be required:

1. A list of the suburbs in Melbourne close to the CBD: this data can be scrapped from Wikipedia ([https://en.wikipedia.org/wiki/List\\_of\\_Melbourne\\_suburbs](https://en.wikipedia.org/wiki/List_of_Melbourne_suburbs)).
2. The Geo-coordinates (latitude, longitude) associated to the Melbourne suburbs that can be obtained using the Geocoder library.
3. The popular venues data of each suburb from Foursquare using RESTful API calls.