

# **CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS**

## **INTRODUCTION**

This is a capstone project for IBM certification course on Data Science. In this project, I am using a hypothetical scenario where a yoga instructor currently residing in Canada wants to find the best place for starting a yoga studio.

In the middle of hustle and bustle of the modern life, emotional stability declines day by day. Yoga is the perfect solution to refurbish attitude towards life by maintaining a positive physical and mental health. With this social purpose in mind, it is also important to find the right location in order to reduce the competition as the yoga instructor is an entrepreneur as well. Hence I am designing this project to help this person to find a best location to start a yoga studio with data analysis skills acquired through the entire span of courses.

## **BUSINESS PROBLEM**

The objective of this capstone project is to find the most suitable location for the entrepreneur to start a new Yoga Studio in Toronto, Canada. By using data science methods and tools along with machine learning algorithms such as clustering, this project aims to provide solutions to answer the business question: In Toronto, if an entrepreneur wants to open a Yoga Studio, where should they consider opening it?

## **INTERESTED AUDIENCE**

The entrepreneur who wants to find the best location to open a Yoga Studio.

## **DATA**

To solve this problem, we will need below data:

- List of neighborhoods in Toronto, Canada
- Latitude and Longitude of these neighborhoods
- Venue data related to existing Yoga Studios. This will help us find neighborhoods that are more suitable to open a Yoga Studio.

## **EXTRACTING THE DATA**

- The scrapping of Toronto neighborhoods via Wikipedia.
- Getting Latitude and Longitude data of these neighborhoods using suitable method (via Geocoder package or csv method).

- Using Foursquare API to get venue data details related to these neighborhoods.