# Applied Data Science Capstone

## IBM Data Science Professional

# Opening a new Bookstore in Toronto, Canada

By: Syed Yusuf

April 2020

## INTRODUCTION

This is a capstone project for IBM Data Science Professional Certificate. In this project, I am creating a hypothetical scenario where a Bookstore wants to expand its franchise in the city of Toronto. Hence the main aim of this project is to help the bookstore decide the best location to start a new outlet of the bookstore.

## BUSINESS PROBLEM

The objective of this capstone project is to find the most suitable location for the bookstore to open a new outlet 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 a Bookstore wants to open a new outlet, where should they consider opening it?

## TARGET AUDIENCE

The bookstore owner who wants to find a new location to open an outlet for his bookstore.

## 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 Bookstores. This will help us to cluster the neighborhoods and find the clusters that are more suitable to open a Bookstore.

## EXTRACTING THE DATA

- Scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods