

IBM Data Science Professional Certificate

Capstone - The Battle of Neighborhoods (Week 1)

INTRODUCTION

This is a capstone project for IBM Data Science Professional Certificate. In this project, we will process a hypothetical scenario for opening a new bookstore in New York City. Bookstores always have an important role for people, especially for educated people and students in large cities like NYC. There are many bookstores in NYC and serious considerations are required to open a new bookstore than it seems. Finding the best location to open a bookstore is one of the most important decisions for investors, and we will complete this project to help stakeholders find the most suitable location.

BUSINESS PROBLEM

The objective of this project is to find the most suitable location for the investor to open a new bookstore in NYC. By using data analysis and tools, and machine learning algorithms such as clustering, this project aims to provide solutions to answer the question; ***“If an entrepreneur wants to open a new bookstore, what would be the best location for it in New York City?”***

TARGET AUDIENCE

The investor who wants to open a new bookstore in New York City.

DATA

To find the best solution, we will need the following data;

- List of neighborhoods in New York City.
- Latitude and longitude coordinates of these neighborhoods.
- Venue data related to bookstores. This will help us find the neighborhoods which are more suitable to open a bookstore.

EXTRACTING THE DATA

- Scrapping data of NYC neighborhoods from https://cocl.us/new_york_dataset
- Getting latitude and longitude data of NYC neighborhoods via Geocoder library
- Using Foursquare API to get venue data related to NYC neighborhoods