

Applied Data Science Capstone Project

1. INTRODUCTION

This is the capstone project for IBM Data Science Professional Certificate. In this one, I am creating a hypothetical scenario that there might not be enough Asian restaurants in New York area. Therefore, it might be a great chance for an entrepreneur who is now looking for an opportunity to start a business in these neighborhoods. This entrepreneur might be thinking of opening his business in areas where Asian community reside. With that purpose in mind, finding an appropriate location to open such a restaurant is one of the most important decisions for this entrepreneur and I am designing this project to help him find the most suitable location.

2. BUSINESS PROBLEM

The objective of this capstone project is to find the most suitable location for the entrepreneur to open a new Asian restaurant in New York City, NY. By using machine learning methods and tools along with appropriate algorithms such as clustering, the goal of this project is to provide solutions to answer the business question: Where should an entrepreneur consider opening an Asian restaurant in New York City?

3. DATA

To solve this problem, we will need below data:

- List of neighborhoods in New York City, NY
- Latitude and Longitude of these neighborhoods
- Venue data related to Asian restaurants. This will help us find the neighborhoods that are more suitable to open an Asian restaurant.

4. PROCESSING THE DATA

- Collecting New York neighborhoods via the file newyork_data.json
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package.
- Using Foursquare API to get venue data related to these neighborhoods.