

Capstone Project - The Battle of Neighborhoods (Week 1)

Finding an optimal location to open a Thai massage & spa in North Jersey

1. Introduction

The project aims to demonstrate how Data Science can leverage decision making in starting a business. Coming up with the question “ If an entrepreneur is trying to open a Thai massage and spa business, where would you recommend him/her? Refer to the question, this section, we will focus on 1) factors that are relevant to selecting a location for the business and 2) an interesting region for the business.

1.1 How to select the right location for Thai massage and spa

First of all, let's get to know Thai massage. Thai massage is a kind of Asian massage that follows the principles of traditional Asian medicine. Thai massage is a full body contact massage that uses a combination of Indian Ayurvedic principles, acupressure massages, and yoga poses. During the massage, the therapist will be following identified lines along the body, and the customer will be positioned in a way that they follow these lines. The customer would feel relieved after the session.

When deciding on a location for the massage and spa, there are a lot of factors to consider—traffic patterns, parking, public transportation, competitors, population density, surrounding businesses, neighborhood ambiance, zoning and rent.

As mentioned above, some of the factors can be revealed by applying location data through Data Science method, for instance, competitors, population density, surrounding businesses, neighborhood ambiance.

1.2 The interesting region

New Jersey is the fourth-smallest state by area but the 11th-most populous, with 8,882,190 residents as of 2019 and an area of 8,722.58 square miles, making it the most densely populated of the 50 U.S. states. The highest population density can reasonably lead to driving traffic to the business and sharing data of venues in cities. In case of the project, these would potentially benefit both the entrepreneur and the use of data analysis.

However, to make the scope of the interesting region smaller, this project will choose a region of New Jersey which is often broadly divided into North Jersey, Central Jersey, and South Jersey. Consequently, North Jersey becomes our choice due to a high level of economic output and its location, connecting to Upper Manhattan in New York City. The following counties are considered North Jersey.

- Bergen County
- Union County
- Essex County
- Hudson County

- Morris County
- Passaic County
- Sussex County
- Warren County

1.3 Project goal

The goal of the project is to recommend an optimal location related to the factors to an entrepreneur planning to open a Thai massage and spa business in the North Jersey region.

2. Data acquisition

The project requires quite a specific dataset of counties, cities and boroughs in New Jersey. This leads to complications of getting data step and most of the available sources have not collected the data up to date. To obtain the latest data of all the counties in North Jersey, 2019 demographic data, such as population, median household income and races from different sources, needs to combine to set up tables. After primarily analyzing the data, the most interesting county for the business is identified. Based on the business problem, next the factors that will influence the decision and how to obtain the data is identified as below:

- The population density and median household income to drive high traffic of the business
- The number of existing competitors in the neighborhood to avoid a high competitive situation.
- The most common venues of the cities and boroughs to estimate surrounding businesses and neighborhood ambiance.

Due to the factors, we need to get more data as following:

- The coordinate of candidate cities and boroughs in the selected county will be obtained using ArcGIS API. The ArcGIS API for Python is a powerful, modern and easy to use Pythonic library to perform GIS visualization and analysis, spatial data management and GIS system administration.
- To create choropleth maps, the GeoJSON files use spatial data is of counties in New Jersey developed by the NJ Office of Information Technology and City of Newark Open Data Portal.
- Foursquare API is used to explore the number of competitors and surrounding businesses in the cities and boroughs.

Add-on tip: New Jersey is divided into 21 counties and contains 565 municipalities consisting of five types: 254 boroughs, 52 cities, 15 towns, 241 townships, and 3 villages. This means in a county can consist of 5 types or less than of the municipalities. Many of the cities were boroughs that sought recognition as cities as they grew bigger. Townships are generally even smaller than boroughs but bigger than town and village.

Source:

<https://vacationidea.com/spas/asian-massage.html>

<https://www.mindbodyonline.com/business/education/blog/how-start-spa-business>

https://en.wikipedia.org/wiki/New_Jersey

<https://developers.arcgis.com/python/guide/overview-of-the-arcgis-api-for-python/>

<https://why.org/articles/explainer-cities-boroughs-and-townships-oh-my-pa-municipalities-clarified/>