Finding the Best Coffee Shop Locations in Boston

Phuong Truong

July 28, 2021

Introduction

Boston is one of the oldest and most densely populated cities in the United States. It has a rich history, vibrant culture and is considered a world leader in higher education. It also attracts millions of tourists, young students and professionals every year. Therefore, Boston is the perfect place for entrepreneurs to start a business or open a restaurant.

This project aims to find an optimal location for a new coffee shop in Boston. Specifically, this project is targeted to stakeholders interested in opening a coffee shop or a café in Boston. Since there are already so many coffee shops in Boston, we need to locate areas that have no or few coffee shops and are densely populated. We also prefer locations that have many universities and young students as these are the main customers for a coffee shop.

Foursquare API and k-means clustering algorithm are used to segment the neighborhoods in Boston. Stakeholders can choose the best neighborhoods for their coffee shops based on the above criteria accordingly.

Data

Based on the definition of the problem, factors that influence the decision are: number of existing café or coffee shops in the neighborhood and population of the neighborhood. In order to segment the neighborhoods and explore them, I need a dataset that contains the neighborhoods as well as the zip code, population and latitude and longitude coordinates of each neighborhood. The following data sources were utilized to extract and generate the required information:

- Boston neighborhoods based on zip code charts from The Boston Globe [1] and Zip Code Business Patterns [2]. Some neighborhoods were removed because they are not located in the City of Boston (Cambridge, Brookline and Somerville). Some zip codes that are attached to a single high-volume address were also removed.

- Massachusetts zip codes by population data from US Census [3].

Example of dataset after extracting:

	ZipCode	Neighborhood	Population
0	02108	Beacon Hill	4454
1	02116	Back Bay	23157
2	02199	Prudential	1157
3	02109	Markets / Inner Harbor	4190
4	02110	Financial District / Wharves	2314

I used the Geocoder Python package to get the geographical coordinates of each neighborhood. [4]

I used Foursquare API to explore common venues in Boston's neighborhoods.

Reference

1. http://archive.boston.com/news/local/articles/2007/04/15/sixfigurezipcodes_city/

2. Boston Redevelopment Authority Research Division, May 2016.

3. https://www.massachusetts-demographics.com/zip_codes_by_population

4. https://geocoder.readthedocs.io/index.html.