Coursera Capstone Project

IBM Applied Data Science Capstone

Picking Neighborhood for Prospective Resident of Boston

By: Chang Liu

January 2020

Introduction

This report is for the IBM Capstone Project of Applied Data Science. A 9-courses series created by IBM, hosted on Coursera platform. The problem and the analysis approach are left for the learner to decide, with a requirement of

leveraging the Foursquare location data to explore or compare neighborhoods or cities of your choice or to come up with a problem that you can use the Foursquare location data to solve.

The main goal will be exploring the neighborhoods of Boston in order to extract the correlation between the average rent and its number of surrounding venues. In other words, a hypothesis that more surrounding venues will lead to a high average rent price for a neighborhood will be tested.

Besides, neighborhoods in Boston will be clustered based on similarities of the most common top 10 venues in each Boston neighborhoods.

The idea comes from the process of a normal family finding a place to stay before moving to another city. It is common that the owners or agents advertise their properties are closed to some kinds of venues like supermarkets, restaurants or coffee shops, etc.; showing the “convenience” of the location in order to raise their house’s value.

So, can the level of “convenience” (the amount of venues in a neighborhood) affect the price of the rent? If so, how strong this relationship between the number of venues and rent price will be?

Target Audiences of this project are listed as below:

1. Prospective residents for Boston who want to pick certain neighborhood for staying with concerns of the conveniences and rent price.
2. Current residents who want to pick a similar neighborhood for moving into a new place.
3. Real estate companies who want to optimize their advertisement and provide suggestions for helping their customers pick a new or a similar neighborhood with the neighborhood where they are living for now.

Business Problem

The objective of this capstone project is to guide potential residents of Boston to pick their most suitable neighborhood before moving in. Using data science methodology and machine learning techniques like clustering, this project aims

to provide solutions to answer the business question: In the city of Boston, MA, US, if a potential resident for Boston is looking for renting a apartment, where should we suggest him/her to rent in terms of the level of “convenience ” of certain neighborhood and how many alternative neighborhoods are there based on the neighborhood groups?