Opening a new Pet Store in Seattle, WA

IBM Applied Data Science Capstone Project
Week 4 Report

By Yiwei Zhang
May 15, 2020

Introduction

Residents in Seattle enjoy their pets' company. According to a survey conducted by Seattletimes, there're even more pets than children in year 2019. With such large market for pets' business, opening a pet store in Seattle seems promising.

A pet store is a retail business which sells variety of animal supplies and pet accessories. These products include: food, treats, toys, collars, leashes, cages and aquariums. Many pet stores also offer pet beauty service and pet training courses. For pet owners, a local pet store brings much convenience to their lives. Yet two similar pet stores that are close to each other may introduce competition. Therefore, the location of a pet store is one of the most important decisions that will determine whether the pet store will be a success or a failure.

Business Problem

The object of this project is to analyze venues for different neighborhoods and select the best locations to open a new pet store in Seattle. By saying the best location, we want the selected area to be home to many residents and has little competition from other local pet stores.

We need to answer the business question: A businessman decides to open a pet store in Seattle, where would you recommend him to open it?

Target Audience

Companies or businessmen that plan to open a new pet store in Seattle is the target audience of this project. Other retail business companies which sell particular kind of items, such as sports shops, shoe shops, can also find the data analysis techniques applied in this project useful for their own business location selection.

Data

To solve this problem, we need the following data:

1. A list of neighborhood names of Seattle.
2. Latitude and longitude data of the above neighborhoods.
3. Venue data of each neighborhood.

Data Sources and Methods to Obtain Data

We can scrap webpage https://en.wikipedia.org/wiki/Category:Neighborhoods_in_Seattle to get a list of neighborhoods of Seattle, with the help of Python package beautifulsoup. After scrapping the raw data, we also need to clean the format of this list which can be done by self-defining python functions. In the end, we will get 27 such neighborhoods.
After getting a formatted list of neighborhood lists, we will use another Python module geocoder to get the geographic coordinates of each neighborhood.
With the longitude and latitude data, we will use Foursquare API to get venue information and explore these neighborhoods.

Data Analysis Techniques to Explore Neighborhoods

We will use k-means clustering algorithm to divide the above 27 neighborhoods into 3 clusters according to the density level of existing pet stores in each neighborhood: barely no pet stores, moderate density, and high density pet stores.
After that, we will use python module folium to visualize the clustering results on a map and provide our suggestion.